

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:02 ; Search time 118.14 Seconds
(without alignments)
235.123 Million cell updates/sec

Title: US-09-596-746a-24

Perfect score: 1842

Sequence: 1 MACNNSGKDCGNTSANSADSE.....KLSKAKEMLTNSVKELTS 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq-1101.*

1:	/SID8/gcgdata/geneseq/AA1980.DAT.*
2:	/SID8/gcgdata/geneseq/AA1981.DAT.*
3:	/SID8/gcgdata/geneseq/AA1982.DAT.*
4:	/SID8/gcgdata/geneseq/AA1983.DAT.*
5:	/SID8/gcgdata/geneseq/AA1984.DAT.*
6:	/SID8/gcgdata/geneseq/AA1985.DAT.*
7:	/SID8/gcgdata/geneseq/AA1986.DAT.*
8:	/SID8/gcgdata/geneseq/AA1987.DAT.*
9:	/SID8/gcgdata/geneseq/AA1988.DAT.*
10:	/SID8/gcgdata/geneseq/AA1989.DAT.*
11:	/SID8/gcgdata/geneseq/AA1990.DAT.*
12:	/SID8/gcgdata/geneseq/AA1991.DAT.*
13:	/SID8/gcgdata/geneseq/AA1992.DAT.*
14:	/SID8/gcgdata/geneseq/AA1993.DAT.*
15:	/SID8/gcgdata/geneseq/AA1994.DAT.*
16:	/SID8/gcgdata/geneseq/AA1995.DAT.*
17:	/SID8/gcgdata/geneseq/AA1996.DAT.*
18:	/SID8/gcgdata/geneseq/AA1997.DAT.*
19:	/SID8/gcgdata/geneseq/AA1998.DAT.*
20:	/SID8/gcgdata/geneseq/AA1999.DAT.*
21:	/SID8/gcgdata/geneseq/AA2000.DAT.*
22:	/SID8/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1837	99.7	374	22	AAB62710
2	1834	99.6	398	22	AAB62728
3	1819.5	98.8	401	22	AAB62738
4	1810.5	98.3	378	22	AAB62712
5	1592.5	86.5	373	22	AAB62711
6	1589.5	86.3	397	22	AAB62729
7	1575	85.3	400	22	AAB62739
8	1566	85.0	377	22	AAB62713
9	1529	83.0	370	22	AAB62715
10	1526	82.8	394	22	AAB62730
11	1493	81.1	368	22	AAB62718

12	1492	81.0	560	22	AAB62724	Borrelia sp chimera
13	1490	80.9	392	22	AAB62734	Borrelia sp chimera
14	1488	80.8	384	22	AAB62726	Borrelia sp chimera
15	1485	80.6	408	22	AAB62737	Borrelia sp chimera
16	1284.5	69.7	369	22	AAB62716	Borrelia sp chimera
17	1281.5	69.6	393	22	AAB62731	Borrelia sp chimera
18	1248.5	67.8	367	22	AAB62719	Borrelia sp chimera
19	1245.5	67.6	391	22	AAB62735	Borrelia sp chimera
20	1213.5	65.9	378	22	AAB62725	Borrelia sp chimera
21	1212.5	65.8	369	22	AAB62714	Borrelia sp chimera
22	1209.5	65.7	393	22	AAB62732	Borrelia sp chimera
23	1205.5	65.4	401	22	AAB62733	Borrelia sp chimera
24	1181.5	64.1	368	22	AAB62717	Borrelia sp chimera
25	1173.5	63.7	391	22	AAB62736	Borrelia sp chimera
26	1165	63.2	410	22	AAB62740	Borrelia sp chimera
27	1164	63.2	386	22	AAB62727	Borrelia sp chimera
28	964	52.3	466	16	AAR5740	B31 outer surface
29	928	50.4	192	22	AAB62721	B burgdorferi ospc
30	928	50.4	210	16	AAW11935	B. burgdorferi str
31	928	50.4	210	16	AAW11935	B. burgdorferi str
32	927	50.3	587	16	AAR5746	B31 Osp-A/antigen
33	926	50.3	192	15	AAR60886	B burgdorferi ospc
34	920	49.9	209	22	AAB62720	B31 outer surface
35	915	49.7	466	16	AAR5739	B burgdorferi ospc
36	908	49.3	192	22	AAB62703	B burgdorferi ospc
37	885	48.0	190	22	AAB62707	B burgdorferi ospc
38	879	47.7	192	15	AAR60889	Borrelia 297 anti
39	864	46.9	176	15	AAR62772	Borrelia B31 anti
40	852	46.3	176	15	AAR62775	Borrelia 297 anti
41	694.5	37.7	177	15	AAR62773	Borrelia 25015 ant
42	694.5	37.7	177	15	AAR60887	Borrelia 25015 Osp
43	678.5	36.8	193	22	AAB62709	B burgdorferi ospc
44	676	36.7	212	16	AAW11934	Outer surface prot
45	674.5	36.6	191	15	AAR60897	Borrelia V5461 ant

ALIGNMENTS

RESULT 1

ID AAB62710 standard; Protein: 374 AA.

XX AAB62710;

AC AAB62710;

XX 03-APR-2001 (first entry)

XX

DE Borrelia sp chimera ospc protein SEQ ID NO: 24.

XX

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

XX

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

XX

PN WO200078966-A1.

XX

PD 28-DEC-2000.

XX

PF 19-JUN-2000; 2000WO-US16915.

XX

PR 18-JUN-1999; 99US-0140042.

XX

PA (UYNV) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX

PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX

DR WPI: 2001-050113/06.

DR N-PSDB; AAF29014.

XX

PT Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

XX Claim 43; Page 78-79; 160pp; English.
 PS The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 374 AA;

Query Match
 Best Local Similarity 99.7%; Score 1837; DB 22; Length 374;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 61
 2 ACPNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 60
 3 GKKIHQNNGLDTEYNHNGSLAGAVAI STLKOKLDGLKNGEIKEDDAKCKSEFTTNK 121
 4 61 gkkihngldteynhngslagavaislklkgldgkngelkedaakckseftnk 120
 5 122 LKEKHTDLGEGVTDADAEALIKNGTKTGAEELGTFESYEVLSKAEMIANSVKE 181
 6 121 lkehhdldggytdadakealilnglktgaelgktsesvevlskaakemiansvke 180
 7 182 LTPVVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAAVEEALIS 241
 8 181 ltpvvaespamvnnsgkdntsansadsvkgnplteiskittesnavlavkeveall 240
 9 242 TSIDELAKAIGKIKINDVSLDNEADHNGSLISGAVLISMLITRKISAIKDSGELEKAE 301
 10 241 tsidelakaigkikndvslndeadhngslisgavylismilitrksaikdsgelekae 300
 11 302 AKKSEFTAKLGEHNDLGEVTDADNAKKAAILKTNNDKTKGAELEKTFESYKNSLKA 361
 12 301 akkseeflaklgehndlgkevtdadnakkailktndktkgadelektfesvknslka 360
 13 362 AKEMLTNSVKEILTS 375
 14 361 akemltinsvkeilts 374

RESULT 2

562728
 AAB62728 standard; Protein; 398 AA.

AC AAB62728;
 DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 60.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
 XX Chimeric - Borrelia sp.
 PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M.

DR WPI: 2001-050113/06.
 DR N-PSDE; AAF29032.
 XX Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease.
 XX
 PS Claim 43; Page 123; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 398 AA;

Query Match
 Best Local Similarity 99.6%; Score 1834; DB 22; Length 398;
 Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 61
 2 ACPNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAAVEEALISIDEIAAKAI 84
 3 GKKIHQNNGLDTEYNHNGSLAGAVAI STLKOKLDGLKNGEIKEDDAKCKSEFTTNK 121
 4 62 gkkihngldteynhngslagavaislklkgldgkngelkedaakckseftnk 144
 5 85 gkkihngldteynhngslagavaislklkgldgkngelkedaakckseftnk 144
 6 122 LKEKHTDLGEGVTDADAEALIKNGTKTGAEELGTFESYEVLSKAEMIANSVKE 181
 7 121 lkehhdldggytdadakealilnglktgaelgktsesvevlskaakemiansvke 204
 8 182 LTPVVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKITTESNAVLAAVEEALIS 241
 9 181 ltpvvaespamvnnsgkdntsansadsvkgnplteiskittesnavlavkeveall 264
 10 242 TSIDELAKAIGKIKINDVSLDNEADHNGSLISGAVLISMLITRKISAIKDSGELEKAE 301
 11 241 tsidelakaigkikndvslndeadhngslisgavylismilitrksaikdsgelekae 324
 12 302 AKKSEFTAKLGEHNDLGEVTDADNAKKAAILKTNNDKTKGAELEKTFESYKNSLKA 361
 13 325 akkseeflaklgehndlgkevtdadnakkailktndktkgadelektfesvknslka 384
 14 362 AKEMLTNSVKEILTS 375
 15 385 akemltinsvkeilts 398

RESULT 3

562738
 AAB62738 standard; Protein; 401 AA.

AC AAB62738;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 80.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
 XX Chimeric - Borrelia sp.
 PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

```

XX (UNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI: 2001-050113/06.
XX N-PSDB: AAF29042.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 150-151; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
XX chimeric ospC proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
XX Sequence 401 AA:
XX
XX Query Match 98.8%; Score 1819.5; DB 22; Length 401;
XX Best Local Similarity 98.7%; Pred. No. 5.6e-118;
XX Matches 372; Conservative 2; Mismatches 0; Indels 3; Gaps 1;
XX
XX 2 ACNNSGKDCNTSANSADSVKGPNTLEISKRTIDSNVAVLLAVEFALLSSIDEIAKAI 61
XX :|||||
XX 25 scnnsgkdgntsaansadesvkgpnlteiskrtidsnavllavkeveallssideiaakai 84
XX
XX 62 GKRIHONNGLDTEYNNHNSGLAGAVAIISPLIKOKLDLNGEIKERIDAAKCCSEFTNK 121
XX 85 gkllhngldteyhnngslagayaislklqkldglnegikerkidaakkcseftnk 144
XX
XX 122 LKEKHTDLKEGYTODAEAILKTNGTKTGAEEIGKLFESVEVLSKAAKEMLANSVKE 181
XX :|||||
XX 145 lkehtdlygeytdadaeailktngtktgaeelgklfesvevlskaakemlansvke 204
XX
XX 182 LKSPVVAES---PAMVNNSGKDCNTSANSADSVKGPNTLEISKRTITESNAVVLAVKEVE 238
XX :|||||
XX 205 lkspvvaesppkpsmwnsgkdgntsaansadesvkgpnlteiskrtitesnavvlavkeve 264
XX
XX 239 TLLTSIDELAKAIGKKIKNDVSLDNEADHNSGLISGAVYLSNLTIRKISAIKDSGELKAE 298
XX :|||||
XX 265 tlltsidelakalgyklkndvslndeadhngslisgavylsnltirkisaikdsgeklkae 324
XX
XX 299 TEKAKKCEEFPAKLKGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESYKNL 358
XX :|||||
XX 325 tekakkcseefpalklgehtdlygeytdnnaakailktndktgadellekifesyvknl 384
XX
XX 359 SKAAKEMLTNSVKEKLS 375
XX :|||||
XX 385 skaakemltnsvkeklcs 401
XX
XX RESULT 4
XX AAB62712
XX ID AAB62712 standard; Protein: 378 AA.
XX
XX AAB62712;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 28.
XX
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
XX
XX Chimeric - Borrelia sp.
XX
XX WO200078966-A1.

```

```

XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000WO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UNY ) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI: 2001-050113/06.
XX N-PSDB: AAF29016.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX
XX Claim 43; Page 83-84; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
XX chimeric ospC proteins from members of the Borrelia genus. These may be
XX Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
XX Sequence 378 AA:
XX
XX Query Match 98.3%; Score 1810.5; DB 22; Length 378;
XX Best Local Similarity 98.1%; Pred. No. 2.2e-117;
XX Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
XX
XX 1 MACNNSGKDCNTSANSADSVKGPNTLEISKRTIDSNVAVLLAVEFALLSSIDEIAKA 60
XX :|||||
XX 1 macnsgkdgntsaansadesvkgpnlteiskrtidsnavllavkeveallssideiaaka 60
XX
XX 61 IGKRIHONNGLDTEYNNHNSGLAGAVAIISPLIKOKLDLNGEIKERIDAAKCCSEFTN 120
XX :|||||
XX 61 igkllhngldtenhngslagayaislklqkldglnegikerkidaakkcseftn 120
XX
XX 121 KLEKHTDLKEGYTODAEAILKTNGTKTGAEEIGKLFESVEVLSKAAKEMLANSVK 180
XX :|||||
XX 121 klehtdlygeytdadaeailktngtktgaeelgklfesvevlskaakemlansvk 180
XX
XX 181 ELKSPVVAES---PAMVNNSGKDCNTSANSADSVKGPNTLEISKRTITESNAVVLAVKEV 237
XX :|||||
XX 181 elkspvvaesppkpsmwnsgkdgntsaansadesvkgpnlteiskrtitesnavvlavkev 240
XX
XX 238 ETLTSLIDELAKAIGKKIKNDVSLDNEADHNSGLISGAVYLSNLTIRKISAIKDSGELKA 297
XX :|||||
XX 241 etltsidelakalgyklkndvslndeadhngslisgavylsnltirkisaikdsgeklka 300
XX
XX 298 ETEKAKKCEEFPAKLKGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESYKN 357
XX :|||||
XX 301 etekakkcseefpalklgehtdlygeytdnnaakailktndktgadellekifesyvkn 360
XX
XX 358 LSKAAKEMLTNSVKEKLS 375
XX :|||||
XX 361 lskaakemltnsvkeklcs 378
XX
XX RESULT 5
XX AAB62711
XX ID AAB62711 standard; Protein: 373 AA.
XX
XX AAB62711;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 26.

```

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 WPI: 2001-050113/06.
 N-PSDB: AAF29015.
 XX
 XX Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 XX to Lyme disease -
 PS Claim 43; Page 81; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 373 AA:

Query Match 86.5%; Score 1592.5; DB 22; Length 373;
 Best Local Similarity 87.7%; Pred. No. 2.4e-102;
 Matches 327; Conservative 20; Mismatches 25; Indels 1; Gaps 1;
 QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
 DB 1 acnnsqkdgntsansadesvkgpnlteiskitdsnavllavkeveallssideiaakai 60
 QY 62 GKRIHONNGLDTEYNNHNSLAGAVAISTLIKOKLDGKNEGLKEKIDAAKCESETFNK 121
 DB 61 gkrihngldteyhnngslagayaistlikqkldgkneglkekidaaakcsetfntk 120
 QY 122 LKEKHTDLGKEGVTDADAKAELIKTNGTKTKGAELGKLFESVEVLSKAAKEMLANSVKE 181
 DB 121 lkehtldlgkegvtdadaakaellktngtktgaeelgklfesvevlskaakemlansvke 180
 QY 182 LTPVVAESPAMVNNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLLAVKEVEALL 241
 DB 181 ltpvvaespamvnnnsqkdgntsansadesvkgpnlteiskitdsnavllavkeveall 240
 QY 242 TSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYVLSLITKIKSAIKDSELAETIE 300
 DB 241 asidelatkaigkikngdvnsladnhngslisgavylslitkiksaikdseleaeie 300
 QY 301 KAKKSEEFTELKGEHTDLGKEGVTDADAKAELIKTNGTKTKGADELKLFESVKNLSK 360
 DB 301 nakksedefftklegehndlgkegvtdadakaellktngtktgadeleklfesvknlsk 360
 QY 361 AAKEMLTNSVKEL 373
 DB 361 aakemlansvkel 373

RESULT 6
 AAB62729
 ID AAB62729 standard; Protein; 397 AA.
 XX

AC AAB62729;
 XX
 XX 03-APR-2001 (First entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 62.
 XX
 XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 WPI: 2001-050113/06.
 N-PSDB: AAF29033.
 XX
 XX Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 XX to Lyme disease -
 PS Claim 43; Page 125-126; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 397 AA:

Query Match 86.3%; Score 1589.5; DB 22; Length 397;
 Best Local Similarity 87.4%; Pred. No. 4.2e-102;
 Matches 326; Conservative 21; Mismatches 25; Indels 1; Gaps 1;
 QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
 DB 25 acnnsqkdgntsansadesvkgpnlteiskitdsnavllavkeveallssideiaakai 84
 QY 62 GKRIHONNGLDTEYNNHNSLAGAVAISTLIKOKLDGKNEGLKEKIDAAKCESETFNK 121
 DB 61 gkrihngldteyhnngslagayaistlikqkldgkneglkekidaaakcsetfntk 144
 QY 122 LKEKHTDLGKEGVTDADAKAELIKTNGTKTKGAELGKLFESVEVLSKAAKEMLANSVKE 181
 DB 121 lkehtldlgkegvtdadaakaellktngtktgaeelgklfesvevlskaakemlansvke 180
 QY 182 LTPVVAESPAMVNNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLLAVKEVEALL 241
 DB 205 ltpvvaespamvnnnsqkdgntsansadesvkgpnlteiskitdsnavllavkeveall 264
 QY 242 TSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYVLSLITKIKSAIKDSELAETIE 300
 DB 241 asidelatkaigkikngdvnsladnhngslisgavylslitkiksaikdseleaeie 300
 QY 301 KAKKSEEFTELKGEHTDLGKEGVTDADAKAELIKTNGTKTKGADELKLFESVKNLSK 360
 DB 325 nakksedefftklegehndlgkegvtdadakaellktngtktgadeleklfesvknlsk 360
 QY 361 AAKEMLTNSVKEL 373
 DB 385 aakemlansvkel 397


```

RESULT 7
AAB62739
ID AAB62739 standard; Protein; 400 AA.
XX
AC AAB62739;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 82.
XX
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PI 19-JUN-2000; 2000WO-US16915.
PR 18-JUN-1999; 99US-0140042.
PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
DR N-PSDB: AAF29043.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 153; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 400 AA:

Query Match 85.5%; Score 1575; DB 22; Length 400;
Best Local Similarity 86.4%; Pred. No. 4.3e-101;
Matches 325; Conservative 22; Mismatches 25; Indels 4; Gaps 2;

```

```

DB 325 kienakcsedftkklegenhqgienvdenakkallitaaakdgaaelexlfkaven 384
QY 358 LSKAKEMLTNSVKEL 373
DB 385 lakaakemlansvkel 400

RESULT 8
AAB62713
ID AAB62713 standard; Protein; 377 AA.
XX
AC AAB62713;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PI 19-JUN-2000; 2000WO-US16915.
PR 18-JUN-1999; 99US-0140042.
PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
DR N-PSDB: AAF29017.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 86-87; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 377 AA:

Query Match 85.0%; Score 1566; DB 22; Length 377;
Best Local Similarity 85.9%; Pred. No. 1.7e-100;
Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

```


QY 62 GKRIHONNGLDTEYHNHNGSLAGAVIAISTLIKOKLDGLK-NEGCKEKIDAAKCKSETFTN 120
 Db 84 gkikndgsldneanrneesllagaylscilqklakngseglakeiaakckseefst 143
 QY 121 KKEKHNDGKRGVNDADAKKALIKTNGT-KTGAEELGKLFESVYLSAAKEMLANSV 179
 Db 144 ktkdnhaqigvgvtdenakkaillkanaagkdkgveelelelsgleslskaakemlanav 203
 QY 180 KKLTSFVVAESPAMVNSGDKGNTSANSADSEYKGNLTETSKITESNAVLAVERET 239
 Db 204 keltepvng-----nsgkdgntsansadesvkgpnlcelskkitesnavlavrevel 258
 QY 240 LITSIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAET 299
 Db 259 lltsidelakaiqkklkndvslidneadhngslisgayllsnlltkklsaltdsgelkael 318
 QY 300 EKAKCKSEEFKAKGEGHTDLGEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLS 359
 Db 319 ekakckseeflaklqghcdlqgyvtdnakkailtkmndckgadelkltesyknls 378
 QY 360 KAAKEMLTNSVKELTS 375
 Db 379 kaekmltnsvkelts 394

RESULT 11
 AAB62718
 ID AAB62718 standard; Protein: 368 AA.
 AC AAB62718;
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 40.
 DE Borrelia; ospc; Lyme disease; vaccine; chimeric; protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 PD 28-DEC-2000.
 PF 19-JUN-2000; 2000WO-US16915.
 PR 18-JUN-1999; 99US-0140042.
 XX (UYNY) UNITV NEW YORK STATE RES FOUND.
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI: 2001-050113/06.
 DR N-PSDB: AAF29022.
 PT Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 PS Claim 43; Page 99-100; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX Sequence 368 AA;
 SQ

Query Match 81.1%; Score 1493; DB 22; Length 368;

Best Local Similarity 83.4%; Pred. No. 1.8e-95;
 Matches 312; Conservative 22; Mismatches 34; Indels 6; Gaps 2;
 QY 2 ACNNSGDKGNTSANSADSEYKGNLTETSKITDSNAVLLAVKEVALSSIDEIAAKAI 61
 Db 1 acnnsygdgnasansadesvkgpnlteiskkitesnavlavreveltlasidelackel 60
 QY 62 GKRIHONNGLDTEYHNHNGSLAGAVIAISTLIKOKLDGLKNEGKLEKIDAAKCKSETFTN 121
 Db 61 gkkl-gungleangskntslisgayaalsdlakelnvlnkeelkeldtakqcscteflnk 119
 QY 122 LKEKHNDGKRGVNDADAKKALIKTNGT-KTGAEELGKLFESVYLSAAKEMLANSV 181
 Db 120 lkeehavlgldnldtdnagralikhanckdgaaelelklfavenlskaagclknayke 179
 QY 182 LITSFVVAESPAMVNSGDKGNTSANSADSEYKGNLTETSKITESNAVLAVERETL 241
 Db 180 ltsplvng-----nsgkdgntsansadesvkgpnlcelskkitesnavlavrevel 234
 QY 242 TSIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAET 301
 Db 235 tsidelakaiqkklkndvslidneadhngslisgayllsnlltkklsaltdsgelkael 294
 QY 302 AKKCKSEEFKAKGEGHTDLGEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLS 361
 Db 295 akckseeflaklqghcdlqgyvtdnakkailtkmndckgadelkltesyknlska 354
 QY 362 AKEMLTNSVKELTS 375
 Db 355 ekemltnsvkelts 368

RESULT 12
 AAB62724
 ID AAB62724 standard; Protein: 560 AA.
 AC AAB62724;
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 52.
 DE Borrelia; ospc; Lyme disease; vaccine; chimeric; protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 PD 28-DEC-2000.
 PF 19-JUN-2000; 2000WO-US16915.
 PR 18-JUN-1999; 99US-0140042.
 XX (UYNY) UNITV NEW YORK STATE RES FOUND.
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI: 2001-050113/06.
 DR N-PSDB: AAF29028.
 PT Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 PS Claim 43; Page 112-113; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads

Sequence 560 AA;

Query Match	81.0%;	Score 1492;	DB 22;	Length 560;
Best Local Similarity	83.5%;	Pred. No. 3.6e-95;		
Matches 313; Conservative	20;	Mismatches 40;		

conservative	20;	Mismatches	40
--------------	-----	------------	----

Indels 2; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKCPNLTETSKKITDSSNAVLAVEALLSIDETAAKA 6C
 |||||
 1 macnnsqkdqntsansadsvkcpnltetssnaavlaveallssidetaaak 6C
 |||||

QY 61 IGGKIHQNNGLDTEYNNHNSGLAGAYAIISTLIKQKIDGLKNBGLKEKIDAKKCEFTN 12
|||||
Db 61 IGGKIHQNNGLDTEYNNHNSGLAGAYAIISTLIKQKIDGLKNBGLKEKIDAKKCEFTN 12
|||||

121 KLEKHTDLGKEGVTDADAKAAILKTNKTKTGAELGKLFEEVEVLSKAAKEMLANSVK 18

181 ELTSPVAESPAMVNSGKDNT-SANSADESVKGPNIITFSKRTMEGNVIR NIVETZ 18

181 eltsipvaespaamsngkygdgaastn padesa kypnlte lskitdsnafv lavkevet 24

241 |v|s|d|e|l|a|k|k|a|i|g|k|i|d|n|n|g|l|a|a|l|n|g|s|l|a|g|a|y|a|i|s|t|l|e|k|i|s|k|i|n|l|e|e|k|t|e| 300

[illegible]

359 SKAKEM/ITNSVKEL 373
|||: |||||

ESQUEMA

AB02734
D AAB62734 standard; Protein; 392 AA.

03-APR-2001 (first entry)

Borrelia sp chimeric ospC protein SEQ ID NO: 72.

Chimeric - Borrelia sp.
Chimeric vaccine, vaccine; chimeric protein; tick.

WO200078966-A1.

19-JUN-2000: 000000 000000

18-JUN-1999; 99US-0140042.

(BROO-) BROOK BIOTECHNOLOGIES INC.

WPI; 2001-050113/06.

Compositions of OSCP polypeptides from strains of *Paracoccidioides brasiliensis*

Antibodies and detect immune responses to Lyme disease -

PS Claim 43; Page 139-140; 160pp; English.
XX

infection provides compositions comprising ospc proteins and chimeric ospc proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 392 AA;

Query Match	80.98;	Score 1490;	DB 22;	Length 392;
Best Local Similarity	83.28;	Pred. No. 3.1e-95;		
Matches 311;	Conservative 23;	Mismatches 34;	Indels 5;	Crosses 0

```

2 ACNNSGKDGNTSANSADSEYKGPNTLEISKRTDSSNAVLAWKEVEALLSSIDEIAKAI 61
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

62 GKIHÖNGLDTEYHNHNSLAGAVAISTLIKÖKLDGLKNEGKKEKIDAKKSEFETNK 121
1111 1111
2Y 2Y

122 LKEKHTDLEKESVADAPAKKATIKVNCSTKQK... 1117
123 ... 1118
124 ... 1119
125 ... 1120
126 ... 1121
127 ... 1122
128 ... 1123
129 ... 1124
130 ... 1125
131 ... 1126
132 ... 1127
133 ... 1128
134 ... 1129
135 ... 1130
136 ... 1131
137 ... 1132
138 ... 1133
139 ... 1134
140 ... 1135
141 ... 1136
142 ... 1137
143 ... 1138
144 ... 1139
145 ... 1140
146 ... 1141
147 ... 1142
148 ... 1143
149 ... 1144
150 ... 1145
151 ... 1146
152 ... 1147
153 ... 1148
154 ... 1149
155 ... 1150
156 ... 1151
157 ... 1152
158 ... 1153
159 ... 1154
160 ... 1155
161 ... 1156
162 ... 1157
163 ... 1158
164 ... 1159
165 ... 1160
166 ... 1161
167 ... 1162
168 ... 1163
169 ... 1164
170 ... 1165
171 ... 1166
172 ... 1167
173 ... 1168
174 ... 1169
175 ... 1170
176 ... 1171
177 ... 1172
178 ... 1173
179 ... 1174
180 ... 1175
181 ... 1176
182 ... 1177
183 ... 1178
184 ... 1179
185 ... 1180
186 ... 1181
187 ... 1182
188 ... 1183
189 ... 1184
190 ... 1185
191 ... 1186
192 ... 1187
193 ... 1188
194 ... 1189
195 ... 1190
196 ... 1191
197 ... 1192
198 ... 1193
199 ... 1194
200 ... 1195
201 ... 1196
202 ... 1197
203 ... 1198
204 ... 1199
205 ... 1200
206 ... 1201
207 ... 1202
208 ... 1203
209 ... 1204
210 ... 1205
211 ... 1206
212 ... 1207
213 ... 1208
214 ... 1209
215 ... 1210
216 ... 1211
217 ... 1212
218 ... 1213
219 ... 1214
220 ... 1215
221 ... 1216
222 ... 1217
223 ... 1218
224 ... 1219
225 ... 1220
226 ... 1221
227 ... 1222
228 ... 1223
229 ... 1224
230 ... 1225
231 ... 1226
232 ... 1227
233 ... 1228
234 ... 1229
235 ... 1230
236 ... 1231
237 ... 1232
238 ... 1233
239 ... 1234
240 ... 1235
241 ... 1236
242 ... 1237
243 ... 1238
244 ... 1239
245 ... 1240
246 ... 1241
247 ... 1242
248 ... 1243
249 ... 1244
250 ... 1245
251 ... 1246
252 ... 1247
253 ... 1248
254 ... 1249
255 ... 1250
256 ... 1251
257 ... 1252
258 ... 1253
259 ... 1254
260 ... 1255
261 ... 1256
262 ... 1257
263 ... 1258
264 ... 1259
265 ... 1260
266 ... 1261
267 ... 1262
268 ... 1263
269 ... 1264
270 ... 1265
271 ... 1266
272 ... 1267
273 ... 1268
274 ... 1269
275 ... 1270
276 ... 1271
277 ... 1272
278 ... 1273
279 ... 1274
280 ... 1275
281 ... 1276
282 ... 1277
283 ... 1278
284 ... 1279
285 ... 1280
286 ... 1281
287 ... 1282
288 ... 1283
289 ... 1284
290 ... 1285
291 ... 1286
292 ... 1287
293 ... 1288
294 ... 1289
295 ... 1290
296 ... 1291
297 ... 1292
298 ... 1293
299 ... 1294
300 ... 1295
301 ... 1296
302 ... 1297
303 ... 1298
304 ... 1299
305 ... 1300
306 ... 1301
307 ... 1302
308 ... 1303
309 ... 1304
310 ... 1305
311 ... 1306
312 ... 1307
313 ... 1308
314 ... 1309
315 ... 1310
316 ... 1311
317 ... 1312
318 ... 1313
319 ... 1314
320 ... 1315
321 ... 1316
322 ... 1317
323 ... 1318
324 ... 1319
325 ... 1320
326 ... 1321
327 ... 1322
328 ... 1323
329 ... 1324
330 ... 1325
331 ... 1326
332 ... 1327
333 ... 1328
334 ... 1329
335 ... 1330
336 ... 1331
337 ... 1332
338 ... 1333
339 ... 1334
340 ... 1335
341 ... 1336
342 ... 1337
343 ... 1338
344 ... 1339
345 ... 1340
346 ... 1341
347 ... 1342
348 ... 1343
349 ... 1344
350 ... 1345
351 ... 1346
352 ... 1347
353 ... 1348
354 ... 1349
355 ... 1350
356 ... 1351
357 ... 1352
358 ... 1353
359 ... 1354
360 ... 1355
361 ... 1356
362 ... 1357
363 ... 1358
364 ... 1359
365 ... 1360
366 ... 1361
367 ... 1362
368 ... 1363
369 ... 1364
370 ... 1365
371 ... 1366
372 ... 1367
373 ... 1368
374 ... 1369
375 ... 1370
376 ... 1371
377 ... 1372
378 ... 1373
379 ... 1374
380 ... 1375
381 ... 1376
382 ... 1377
383 ... 1378
384 ... 1379
385 ... 1380
386 ... 1381
387 ... 1382
388 ... 1383
389 ... 1384
390 ... 1385
391 ... 1386
392 ... 1387
393 ... 1388
394 ... 1389
395 ... 1390
396 ... 1391
397 ... 1392
398 ... 1393
399 ... 1394
400 ... 1395
401 ... 1396
402 ... 1397
403 ... 1398
404 ... 1399
405 ... 1400
406 ... 1401
407 ... 1402
408 ... 1403
409 ... 1404
410 ... 1405
411 ... 1406
412 ... 1407
413 ... 1408
414 ... 1409
415 ... 1410
416 ... 1411
417 ... 1412
418 ... 1413
419 ... 1414
420 ... 1415
421 ... 1416
422 ... 1417
423 ... 1418
424 ... 1419
425 ... 1420
426 ... 1421
427 ... 1422
428 ... 1423
429 ... 1424
430 ... 1425
431 ... 1426
432 ... 1427
433 ... 1428
434 ... 1429
435 ... 1430
436 ... 1431
437 ... 1432
438 ... 1433
439 ... 1434
440 ... 1435
441 ... 1436
442 ... 1437
443 ... 1438
444 ... 1439
445 ... 1440
446 ... 1441
447 ... 1442
448 ... 1443
449 ... 1444
450 ... 1445
451 ... 1446
452 ... 1447
453 ... 1448
454 ... 1449
455 ... 1450
456 ... 1451
457 ... 1452
458 ... 1453
459 ... 1454
460 ... 1455
461 ... 1456
462 ... 1457
463 ... 1458
464 ... 1459
465 ... 1460
466 ... 1461
467 ... 1462
468 ... 1463
469 ... 1464
470 ... 1465
471 ... 1466
472 ... 1467
473 ... 1468
474 ... 1469
475 ... 1470
476 ... 1471
477 ... 1472
478 ... 1473
479 ... 1474
480 ... 1475
481 ... 1476
482 ... 1477
483 ... 1478
484 ... 1479
485 ... 1480
486 ... 1481
487 ... 1482
488 ... 1483
489 ... 1484
490 ... 1485
491 ... 1486
492

144 Iksehavylgldnltddnagraalkkhanckkgaaeleikfavenlsksaqtclkaavke 203

```

||||: |||||
204 Itsp1vhg-----msgkqdnatsansadesvknmlfslsklrfornoumlah|||||
||||: |||||

```

242 TSIDELAKAIKKIKNDVSLNEADHNGSLISGAYLLSNLITKISAIKDSGELKAEIER 301

302 AKKSEETAKLGEHTDLGEGVDDNAKAIKLTNNDKTGADELEKLFESYKYLSPA 361

AKEMLTNSYKELTSS 375
AKEMLTNSYKELTSS 375
AKEMLTNSYKELTSS 375

379 akennlmsvkelts 392

SOLIT 14
AB62726

AAB62726;

05-APR-2001 (first entry)

Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

Chimeric - Borrelia sp.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

(U)NY) UNIV NEW YORK STATE RES FOUND.
(BROO-) BROOK BIODIVERSITY

Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M:

WPI; 2001-050113/06.

DR N-PSDB; AAF29030.
 XX Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 XX Claim 43; Page 117-118; 160pp; English.
 PS
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 XX Sequence 384 AA:
 SQ
 Query Match 80.8%; Score 1488; DB 22; Length 384;
 Best Local Similarity 82.0%; Pred. No. 4.1e-95;
 Matches 309; Conservative 31; Mismatches 31; Indels 6; Gaps 3;
 2 ACNNSGKDGNTSANSADSEYKGNLTREISKRTDSNAVLLAVKEVALLSSIDEIAAKAI 61
 Db 1 acnnsqkdgntsansadesykgpnlteiskrtidsnavllavkevallssideiaakai 60
 QY 62 GKRIHONNGIDTEYNHNGSLAGAVAIISTLIKOKLDKNEGLKEKRIIDAKKCEPTNK 121
 Db 61 gkrihngngldteynhngslagayaistlikqkloglkneglkekiidaakkcseftnk 120
 QY 122 LKEKHHDLCGEVTDADAKAEILKTNGTKTGAELGKLFESYEVLSKAKEMLANSVKE 181
 Db 121 lkekhhdldgkgytdadakeailktngtktgaeelgklfesvevlskaakemlansvke 180
 QY 182 LTPVVAESPAMV--NNSGKDGNTSANSADSEYKGNLTREISKRTESNAVLLAVKEVE 238
 Db 181 ltpvvaespdkpfbhngnsq--gdsastnpdesakgpnlteiskrtidsnavllavkeve 238
 QY 239 TLTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYILSNLTITKRIISAKISGELKAE 298
 Db 239 allssideiskalgkrikndvslneadnngslisgayeisltlqglsvl-nseelkxx 297
 QY 299 IERAKKCEEFYAKLKGHEHTDLGEGVTDNNAKKAILEKTNNDKTKGADELEKLFESYKNL 358
 Db 298 lkeakdcsgkftclklkdsnaelqigvqddnakkailktngtkdkgaeleefksiesl 357
 QY 359 SKAAKEMLTNSVKELTS 375
 Db 358 skaagaaltnsvkeltn 374
 DB 358 skaagaaltnsvkeltn 374
 ULT 15
 AAB62737
 ID AAB62737 standard; Protein; 408 AA.
 XX
 AC AAB62737;
 XX
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 78.
 XX
 KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000MO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX

PA (UNYV) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datkwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 XX
 DR WPI; 2001-050113/06.
 DR N-PSDB; AAF29041.
 XX
 PS Claim 43; Page 147-148; 160pp; English.
 CC
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 XX Sequence 408 AA:
 SQ
 Query Match 80.6%; Score 1485; DB 22; Length 408;
 Best Local Similarity 81.7%; Pred. No. 7.2e-95;
 Matches 308; Conservative 32; Mismatches 31; Indels 6; Gaps 3;
 2 ACNNSGKDGNTSANSADSEYKGNLTREISKRTDSNAVLLAVKEVALLSSIDEIAAKAI 61
 Db 25 scnsgkdgntsansadesykgpnlteiskrtidsnavllavkevallssideiaakai 84
 QY 62 GKRIHONNGIDTEYNHNGSLAGAVAIISTLIKOKLDKNEGLKEKRIIDAKKCEPTNK 121
 Db 61 gkrihngngldteynhngslagayaistlikqkloglkneglkekiidaakkcseftnk 144
 QY 85 gkrihngngldteynhngslagayaistlikqkloglkneglkekiidaakkcseftnk 144
 Db 122 LKEKHHDLCGEVTDADAKAEILKTNGTKTGAELGKLFESYEVLSKAKEMLANSVKE 181
 Db 145 lkekhhdldgkgytdadakeailktngtktgaeelgklfesvevlskaakemlansvke 204
 QY 182 LTPVVAESPAMV--NNSGKDGNTSANSADSEYKGNLTREISKRTESNAVLLAVKEVE 238
 Db 205 ltpvvaespdkpfbhngnsq--gdsastnpdesakgpnlteiskrtidsnavllavkeve 262
 QY 239 TLTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYILSNLTITKRIISAKISGELKAE 298
 Db 263 allssideiskalgkrikndvslneadnngslisgayeisltlqglsvl-nseelkxx 321
 QY 299 IERAKKCEEFYAKLKGHEHTDLGEGVTDNNAKKAILEKTNNDKTKGADELEKLFESYKNL 358
 Db 322 lkeakdcsgkftclklkdsnaelqigvqddnakkailktngtkdkgaeleefksiesl 381
 QY 359 SKAAKEMLTNSVKELTS 375
 Db 382 skaagaaltnsvkeltn 398

Search completed: March 18, 2002, 09:54:29
 Job time: 327 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:47 ; Search time 55.5 seconds
(without alignments)
152,049 Million cell updates/sec

Title: US-09-596-746A-24

Sequence: 1 MACNNSGKCDNTSANSADSE.....KNTSKAKEMUTNSVKELTS 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	964	52.3	466	4	US-08-235-836C-110 Sequence 110, App
2	932	50.6	210	4	US-08-235-836C-30 Sequence 30, Appl
3	928	50.4	209	4	US-09-196-293-15 Sequence 15, Appl
4	928	50.4	210	1	US-08-158-353-3 Sequence 3, Appl
5	927	50.3	588	4	US-08-235-836C-122 Sequence 122, App
6	921	50.0	210	4	US-08-209-603E-15 Sequence 15, Appl
7	915	49.7	466	4	US-08-235-836C-107 Sequence 107, App
8	676	36.7	212	1	US-08-031-295-2 Sequence 2, Appl
9	661	35.9	212	1	US-07-903-580-2 Sequence 2, Appl
10	661	35.9	212	1	US-07-903-580-2 Sequence 2, Appl
11	614	33.3	212	1	US-08-158-353-4 Sequence 4, Appl
12	612	33.2	212	4	US-09-196-293-11 Sequence 11, Appl
13	612	33.2	212	4	US-08-209-603E-11 Sequence 11, Appl
14	609	33.1	212	4	US-08-235-836C-34 Sequence 34, Appl
15	600	32.6	209	4	US-08-235-836C-32 Sequence 32, Appl
16	583.5	31.7	207	4	US-08-235-836C-36 Sequence 36, Appl
17	163.5	8.9	248	1	US-08-328-234-6 Sequence 6, Appl
18	162.5	8.8	3248	5	US-08-353-700-1 Sequence 1, Appl
19	162.5	8.8	3248	5	US-08-353-700-1 Sequence 1, Appl
20	157	8.5	194	4	US-09-364-083-2 Sequence 2, Appl
21	148.5	8.1	1713	3	US-08-600-982-24 Sequence 24, Appl
22	148.5	8.1	1713	3	US-08-600-982-24 Sequence 24, Appl
23	147	8.0	1164	4	US-08-923-992A-2 Sequence 2, Appl
24	146	7.9	1196	1	US-08-144-121-4 Sequence 4, Appl
25	146	7.9	1196	2	US-08-735-893-4 Sequence 4, Appl
26	145.5	7.9	1388	2	US-08-685-576-1 Sequence 1, Appl
27	145	7.9	1588	5	PCT-US93-07261-11 Sequence 11, Appl

28	145	7.9	1663	5	PCT-US93-07261-16 Sequence 16, Appl
29	145	7.9	3111	2	US-08-460-309-4 Sequence 4, Appl
30	145	7.9	3111	2	US-08-125-077-4 Sequence 4, Appl
31	143.5	7.8	630	4	US-08-973-462-9 Sequence 9, Appl
32	143	7.8	1164	4	US-08-923-992A-10 Sequence 10, Appl
33	142.5	7.7	688	3	US-09-141-047-8 Sequence 8, Appl
34	141	7.7	1786	4	US-08-973-462-8 Sequence 8, Appl
35	139	7.5	2285	4	US-09-308-375-2 Sequence 2, Appl
36	138.5	7.5	1098	4	US-08-923-992A-8 Sequence 8, Appl
37	138	7.5	1128	4	US-08-923-992A-6 Sequence 6, Appl
38	137	7.4	1104	4	US-08-923-992A-4 Sequence 4, Appl
39	136	7.4	1388	2	US-08-685-576-4 Sequence 4, Appl
40	135.5	7.4	573	4	US-08-235-836C-112 Sequence 112, App
41	133.5	7.2	641	4	US-08-961-083-160 Sequence 160, App
42	132.5	7.2	1561	3	US-08-894-017-23 Sequence 23, Appl
43	131.5	7.1	1068	4	US-09-085-199B-11 Sequence 11, Appl
44	131	7.1	1093	5	PCT-US93-03077-1 Sequence 1, Appl
45	130.5	7.1	396	1	US-08-430-024-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric proteins Comprising
TITLE OF INVENTION: Botrellia Polypeptides and Uses therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESSES:
ADDRESS: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110

Query Match 52.3%; Score 964; DB 4; Length 466;
Best local Similarity 59.6%; Pred. No. 7e-62;
Matches 229; Conservative 31; Mismatches 92; Indels 32; Gaps 9;

QY 1 MACNNSGKCDNTSANSADSEVKPNLIEISKRTTDSNAVLAVKVEALLSIDELAAKA 60

```

Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRITDSNVLAVKEVEALLSSIDEIATAKA 76
Qy 61 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 120
Db 77 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 136
Qy 121 KKEKHTDGLKGEVTDADAKAEAILKTNGTKGAEEIGKLFESVEVLSKAAKEMLANSVK 180
Db 137 KKAHHTDGLKGEVTDADAKAEAILKTNGTKGAEEIGKLFESVEVLSKAAKEMLANSVK 196
Qy 161 ELTSPVAESP--AMVNSGKDGNTSANSADSVKGNPLTEISKRITDSNA--VVLAVKE 236
Db 197 ELTSPVAESP--AMVNSGKDGNTSANSADSVKGNPLTEISKRITDSNA--VVLAVKE 253
Qy 237 VETLLTS-----IDELAKAIGKIKNDVSLDNEADHNSLLSGAVLT--SNLITKIS 287
Db 254 LELGTSQKNNNGSGVLEGVKADSKVLTIS-----DGLQOTLLEVKEDGKTLVSKRYT 308
Qy 288 AKKSGELKAEIEKAKKCEFTAKLKGHTDGLKGEVTDADAKAEAILKTNNDKTGAD 347
Db 309 S-KKSSSTEEKFNKGEVSEKIIIRADG--TRLEYTGIRKDGSGKAKKRYLVLEGTLT 365
Qy 348 LEKLFESY-----KNLSKAAK 363
Db 366 AEKTLTVKEGVTYLSKNISKSGE 389

```

```

RESULT 2
US-08-235-836C-30
; Sequence 30, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-836C-30

```

```

Query Match 50.4%; Score 928; DB 4; Length 209;
Best Local Similarity 98.4%; Pred. No. 4.9e-60;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRITDSNVLAVKEVEALLSSIDEIATAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRITDSNVLAVKEVEALLSSIDEIATAKA 76
Qy 61 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 120
Db 77 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 136
Qy 121 KKEKHTDGLKGEVTDADAKAEAILKTNGTKGAEEIGKLFESVEVLSKAAKEMLANSVK 180
Db 137 KKEKHTDGLKGEVTDADAKAEAILKTNGTKGAEEIGKLFESVEVLSKAAKEMLANSVK 196
Qy 181 ELTSPVAESP 191
Db 197 ELTSPVAESP 207

```

```

RESULT 3
US-09-196-293-15
; Sequence 15, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; APPLICANT: Fuchs, Renate
; APPLICANT: Motz, Manfred
; APPLICANT: Wilschke, Erwin
; APPLICANT: Wilschke, Bettina
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrelia
; FILE REFERENCE: burgdorferi
; CURRENT APPLICATION NUMBER: US/09/196,293
; EARLIER FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1989-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; TYPE: PRF
; LENGTH: 209
; ORGANISM: Borrelia burgdorferi
; US-09-196-293-15

```

```

Query Match 50.4%; Score 928; DB 4; Length 209;
Best Local Similarity 98.4%; Pred. No. 9.5e-60;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRITDSNVLAVKEVEALLSSIDEIATAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRITDSNVLAVKEVEALLSSIDEIATAKA 76
Qy 61 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 120
Db 77 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKQKLDGLKNEGLKEKIDAAKCCSEFTFN 136
Qy 121 KKEKHTDGLKGEVTDADAKAEAILKTNGTKGAEEIGKLFESVEVLSKAAKEMLANSVK 180
Db 137 KKEKHTDGLKGEVTDADAKAEAILKTNGTKGAEEIGKLFESVEVLSKAAKEMLANSVK 196
Qy 181 ELTSPVAESP 191

```


Db 197 ELTSPVAESP 207

RESULT 4

US-08-158-353-3
; Sequence 3, Application US/08158353
; Patent No. 5620862

GENERAL INFORMATION:

APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme

TITLE OF INVENTION: Disease

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Alice O.

REGISTRATION NUMBER: 33,542

REFERENCE/DOCKET NUMBER: UCT93-05

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 210 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-158-353-3

Query Match 50.4%; Score 928; DB 1; Length 210;
Best Local Similarity 98.4%; Pred. No. 9.5e-60;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 MACNNGSGKGNSTANSANDESVKGNPNTETSKTTDSNAVLAVKEVEALLSIDETAAKA 60

17 ISGNNGSGKGNSTANSANDESVKGNPNTETSKTTDSNAVLAVKEVEALLSIDETAAKA 76

61 IGRKIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFTN 120

77 IGRKIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFTN 136

121 KLEKHTDGLKEGVTADAKAEALIKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

137 KLEKHTDGLKEGVTADAKAEALIKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

181 ELTSPVAESP 191

197 ELTSPVAESP 207

RESULT 5

US-08-235-836C-122
; Sequence 122, Application US/08235836C
; Patent No. 6248362

GENERAL INFORMATION:

APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/235,836C

FILING DATE: 29-Apr-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191

FILING DATE: 01-11-93

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 122:

SEQUENCE CHARACTERISTICS:

LENGTH: 588 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-122

Query Match 50.3%; Score 927; DB 4; Length 588;
Best Local Similarity 88.9%; Pred. No. 4.3e-59;
Matches 193; Conservative 4; Mismatches 10; Indels 10; Gaps 2;

4 NNSGKGNSTANSANDESVKGNPNTETSKTTDSNAVLAVKEVEALLSIDETAAKAIGK 63

286 NNSGKGNSTANSANDESVKGNPNTETSKTTDSNAVLAVKEVEALLSIDETAAKAIGK 345

64 KIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFTNKLK 123

346 KIQNNGLDTEVYHNHNSILAGAYAISTLIKOKLDGLKNEGLKREKIDAKKCEFTFTNKLK 405

124 EKHTDGLKEGVTADAKAEALIKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183

406 EKHTDGLKEGVTADAKAEALIKTNGTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 465

184 SPVVAESP-----AMVNN-----SGKGNSTANSAD 210

466 SPVVAESPKKPGTMAQTNMHLNKSASQNVRTAE 502

RESULT 6

US-08-209-603E-15
; Sequence 15, Application US/08209603E
; Patent No. 6248338

GENERAL INFORMATION:

APPLICANT: FUCHS, RENATE

APPLICANT: WILKE, BETTINA

APPLICANT: PREAC-MURSIC, VERA

APPLICANT: MORTZ, MANFRED

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
MOLECULE TYPE: LINEAR

DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
AMINO-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:

ORGANISM: B. BURGDOFFERI
IMMEDIATE SOURCE:

LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:

IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A

US-08-209-603E-15

Query Match

Best Local Similarity 50.0%; Score 921; DB 4; Length 210;
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLAVKEVEALLSSIDEIAKA 60
DB :|||||
QY 61 IGGKHONNGLDTEYNNHNSLLAGAYAISTLIKOKLDGKNEGLKEKIDAAKCCSEFTFN 76
DB :|||||
QY 77 IGGKHONNGLDTEYNNHNSLLAGAYAISTLIKOKLDGKNEGLKEKIDAAKCCSEFTFN 120
DB :|||||
QY 121 KIKKHTDGLKGEVYDADAKKAILKTNGTKTGKGAELKFESEVEVLSKAKEMLANSVK 180
DB :|||||
QY 137 KIKKHTDGLKGEVYDADAKKAILKTNGTKTGKGAELKFESEVEVLSKAKEMLANSVK 196
QY 181 ELTSPVVAESP 191
DB :|||||
QY 197 ELTSPVVAESP 207

RESULT 7
US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:

APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A

TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
MOLECULE TYPE: linear

US-08-235-836C-107

Query Match

Best Local Similarity 49.7%; Score 915; DB 4; Length 466;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 NNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLAVKEVEALLSSIDEIAKA 63
DB :|||||
QY 64 KIHONNGLDTEYNNHNSLLAGAYAISTLIKOKLDGKNEGLKEKIDAAKCCSEFTFN 123
DB :|||||
QY 124 EKHTDGLKGEVYDADAKKAILKTNGTKTGKGAELKFESEVEVLSKAKEMLANSVK 183
DB :|||||
QY 396 AKHTDGLKGEVYDADAKKAILKTNGTKTGKGAELKFESEVEVLSKAKEMLANSVK 455
QY 184 SPVVAESP 191
DB :|||||
QY 456 SPVVAESP 463

RESULT 8

US-08-158-353-2
Sequence 2, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

```

: STREET: Two Millite Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/158,353
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Carroll, Alice O.
: REGISTRATION NUMBER: 33,542
: REFERENCE/DOCKET NUMBER: UCT93-05
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 617-861-9540
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 212 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-158-353-2

```

```

Query Match      36.7%; Score 676; DB 1; Length 212;
Best Local Similarity 74.7%; Pred. No. 1.2e-41;
Matches 145; Conservative 18; Mismatches 27; Indels 4; Gaps 3;

```

```

QY 1 MACNNSGKDONTANSADSVKGNLTETSKITDSNAVLAVKEVEALLSIDEIATAKA 60
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 ISCNNSGKDGSTANSADSVKGNLTETSKITDSNAVLAVKEVEALLSIDEIATAKA 76
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 IGKKIHQNNGLDTEYHNHNSLLAGAVAI STLTKOKLDGLKN-BGLKEKIDAAKCCSEFTT 119
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 IGNLIAQ-NGLNAGANGNSLLAGAVI STLTKOKLDGLKN-BGLKEKIDAAKCCSEFTT 135
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 NKLEKHTDGG--KEGTVDADAEALITKGTGTGAEELGKLFESVEVLSKAKEMLAN 177
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 136 DKKSSHAELGIANGATDANAKAAILKTGTGKGAQLEKLFESVKNLSKAAQETLNN 195
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 SVKELTSPVAESP 191
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 196 SVKELTSPVAESP 209
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9
: US-08-031-295-2
: Sequence 2, Application US/08031295
: Patent No. 5530103
: GENERAL INFORMATION:
: APPLICANT: LIVEY, Ian
: APPLICANT: DORNER, Friedrich
: TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
: TITLE OF INVENTION: PREVENTION OF LYME DISEASE
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/031,295
: FILING DATE: 19930312
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/903,580
: FILING DATE: 25-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/824,161
: FILING DATE: 22-JAN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/727,245
: FILING DATE: 11-JUL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/142 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEEX: 904136
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 212 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-031-295-2

```

```

Query Match      35.9%; Score 661; DB 1; Length 212;
Best Local Similarity 73.6%; Pred. No. 1.5e-40;
Matches 142; Conservative 14; Mismatches 35; Indels 2; Gaps 2;

```

```

QY 1 MACNNSGKDONT-SANSADSVKGNLTETSKITDSNAVLAVKEVEALLSIDEIATAKA 59
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 ISCNNSGKDGSTANSADSVKGNLTETSKITDSNAVLAVKEVEALLSIDEIATAKA 76
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 AIGKKIHQNNGLDTEYHNHNSLLAGAVAI STLTKOKLDGLKN-BGLKEKIDAAKCCSEFTT 118
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 AIGKKIHQNNGLDTEYHNHNSLLAGAVAI STLTKOKLDGLKN-BGLKEKIDAAKCCSEFTT 136
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 TNKLEKHTDGG--KEGTVDADAEALITKGTGTGAEELGKLFESVEVLSKAKEMLAN 178
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 137 TKLIAAGHAGIDGATDANSKAEALITKGTGTGAEELGKLFESVKNLSKAAQETLNN 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 SVKELTSPVAESP 191
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 SVKELTSPVAESP 209
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 10
: US-07-903-580-2
: Sequence 2, Application US/07903580
: Patent No. 6221363
: GENERAL INFORMATION:
: APPLICANT: LIVEY, Ian
: APPLICANT: DORNER, Friedrich
: TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
: TITLE OF INVENTION: PREVENTION OF LYME DISEASE
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25

```

```

/ ATTORNEY/AGENT INFORMATION:
/ NAME: Carroll, Alice O.
/ REGISTRATION NUMBER: 33,542
/ REFERENCE/DOCKET NUMBER: UCT93-05
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-861-6240
/ TELEFAX: 617-861-9540
/ INFORMATION FOR SEO ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 212 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-158-353-4

Query Match          33.3%; Score 614; DB 1; Length 212;
Best Local Similarity 69.9%; Pred. No. 3,5e-37;
Matches 135; Conservative 14; Mismatches 42; Indels 2; Gaps 2

Oy      1  MÄCNSGKQONT-SANSADESYKGNLTLEISKRTDTSNAVLLAVKEVALLSIDETIAK 59
Db      17  ISCNHGSGKGDSTNPNADESAKGNTLLEISKRTDTSNAFVLLAVKEVETLVLSIDETIAK 76

```

[illegible]

178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689

```

12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
```

Best Local Similarity 70.5%; Pred. NO. 4,8e-37;
Matches 136; Conservative 12; Mismatches 43; Indels 2; Gaps 2

OY	MACNNSKCDG- <u>MANSANADESVKPNLTELTSKTTDSNAVLAVKEEALLSIDELAAK</u>	59
Dd	17 ISCNNSGRVGLTSTINPADESAKNPLTELSKTTDSNAVLAVKEETLVLSIDELAAK	76
OY	60 AIGKKIHONNGLDYEYNHNSLLAGAYASTLKOKIDGLKN-EGLEKEIDAARKCSETF	118
Dd	77 AIGKIDNNNLALANNONGSLLAGAYASTLTETKSLKLNDELKTEITAAKAKCSEEF	136
OY	119 TNKLAKERTDLAGKGVDDADAKEILKTNGTKRKGAELOGLRESVVSLSAAREMIANS	178
Dd	137 TNKLKSGIADGGKODADDNDHKAAILKTHATDGAKAEFYDFESVEGLGLKAAOVALTNS	186
OY	179 VKELTSPVAESP 191	
Dd	197 VKELTSPVAESP 209	

RESULT 13
08-209-

Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: MILSKRE, BETTINA
APPLICANT: PREAC-MURISIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUSCIECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

```

: IMMEDIATE SOURCE:
: LIBRARY: DSM 5662
: POSITION IN GENOME: N/A
: FEATURE:
: IDENTIFICATION METHOD: amino acid analysis
: PUBLICATION INFORMATION: N/A
: ANTI-SENSE: N/A
: FRAGMENT TYPE: INTERNAL
: ORIGINAL SOURCE:
US-08-209-603E-11

```

Query Match	33.2%;	Score 612;	DB 4;	Length 212;
Best Local Similarity	70.5%;	Pred. No. 4.8e-37;		
Matches 136; Conservative	12;	Mismatches 43;	Indels 2;	Gaps 2

Qy	60	AIGKIHONNGLDIEVHNHNSLLAGAVASITLIKQIKDGLKN--EGJKEKIDAAKCKSETF	118
Dd	17	ISCNSSGSKVGILTSTNPADESAQNPLETSKTKTTDSNAEVLWKEVEYETLVLSIDEIAKK	76
Qy	60	AIGKIHONNGLDIEVHNHNSLLAGAVASITLIKQIKDGLKN--EGJKEKIDAAKCKSETF	118
Dd	77	AIGKIDINNNGLALNNONSLLAGAASITLTKEKISKLNNEELTELETAARCKSEEF	136
Qy	119	TNKLKEKHTDLGRGVTDADAKKEAIIKYTGNGTKTGAEEELGFESYEVSLSKAACKEMLANS	178
Dd	137	TNKLKSCHADGKODATDDHAKKAAILKTHTATTDKGAKEFDFLESVEGLLKAAOVALTNS	196
Qy	179	VKELTSPVAESP	191
Dd	197	VKELTSPVAESP	209

RESULT 14
US-08-235-

; Sequence 34, Application US/08235836C
; Patent No. 6248562
; CURRENT INFORMATION.

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Lutfi, Benjamin J.
TITLE OF INVENTION: No. 6248562a1 Chimeric Proteins Comprising
TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-34

Query Match 33.1%; Score 609; DB 4; Length 212;
Best Local Similarity 69.4%; Pred. No. 7.9e-37;

Matches 134; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNSGKDGNT-SANSADESVKGPNTLEISKITDSNAVLAVKEVEALLSSIDEIAR 59
Db 17 ISCNNSGKGDSASTNPDESAGPNLTLEISKITDSNAVLAVKEVEALLSSIDEIAR 76
QY 60 AIGKIHQNNGLDTEYVNHGSLLAGAVAISTLIKOKIDGLKN-EGLKEKIDAAKCSFT 118
Db 77 AIGKIDNNNGLAALNNQNSLAGAVAISTLIKTEKSLKLNLEKTEIAKAKCSEEF 136
119 TNKKEKHTDLGKEGYTDADAKAEILKTNGTKTGAEEELKFESEVLSKAKEMLANS 178
137 TNKLSGHDGLGQDADDDHAKAAILKTHATTDGAKAEFDLFESEVGLKAAQVALTNS 196
QY 179 VKELTSPVAESP 191
Db 197 VKELTSPVAESP 209

RESULT 15

US-08-235-836C-32
Sequence 32, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:

APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
NUMBER OF INVENTIONS: Borrellia Polypeptides and Uses therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:

ADDRESS: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235, 836C
FILING DATE: 29-APR-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93

ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 209 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-32

Matches 131; Conservative 19; Mismatches 39; Indels 3; Gaps 2;

QY 1 MACNSGKDGNTSANSADESVKGPNTLEISKITDSNAVLAVKEVEALLSSIDEIAR 60
Db 17 ISCNNSG--GDTASTNPDESAGPNLTLEISKITDSNAVLAVKEVEALLSSIDEIAR 74
QY 61 ICKKIHQNNGLDTEYVNHGSLLAGAVAISTLIKOKIDGLKN-EGLKEKIDAAKCSFT 119
Db 75 ICKVHQNNGNLNANMGQNSLAGAVAISTLIKTEKSLKLNSEELNKITEAKNHSEAF 134
QY 120 NKLEKHTDLGKEGYTDADAKAEILKTNGTKTGAEEELKFESEVLSKAKEMLANS 179
Db 135 NKLKSHQGLGVAATDADHAKAAILKSNPTKDGAKALKDSESEVLSKAAQVALANSV 194
QY 180 KELTSPVAESP 191
Db 195 KELTSPVAESP 206

Search completed: March 18, 2002, 09:55:32
Job time: 345 sec

Query Match 32.6%; Score 600.5; DB 4; Length 209;
Best Local Similarity 68.2%; Pred. No. 3.2e-36;

THIS PAGE BLANK

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:32 ; Search time 621.2 Seconds
(without alignments)
167.613 Million cell updates/sec

Title: US-09-596-746A-24

Sequence: 1 MACNNSGKDGNTSANSADSE.....KNLSKAKEMLTNSYKELTS 375

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 segs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Pending Patents, AA, Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1842	100.0	375	19	US-09-596-746A-24
2	1837	99.7	374	19	US-09-596-746-24
3	1834	99.6	398	19	US-09-596-746-60
4	1834	99.6	399	19	US-09-596-746A-60
5	1819.5	98.8	401	19	US-09-596-746-80
6	1819.5	98.8	402	19	US-09-596-746A-80
7	1810.5	98.3	378	19	US-09-596-746-28
8	1810.5	98.3	378	19	US-09-596-746A-28
9	1597.5	86.7	374	19	US-09-596-746A-26

10	1592.5	86.5	373	19	US-09-596-746-26	Sequence 26, Appl
11	1589.5	86.3	397	19	US-09-596-746-62	Sequence 62, Appl
12	1589.5	86.3	398	19	US-09-596-746A-62	Sequence 82, Appl
13	1575	85.5	400	19	US-09-596-746-82	Sequence 82, Appl
14	1575	85.5	401	19	US-09-596-746A-82	Sequence 82, Appl
15	1566	85.0	377	19	US-09-596-746-30	Sequence 30, Appl
16	1566	85.0	377	19	US-09-596-746A-30	Sequence 30, Appl
17	1534	83.3	371	19	US-09-596-746A-30	Sequence 30, Appl
18	1529	83.0	370	19	US-09-596-746A-34	Sequence 34, Appl
19	1526	82.8	394	19	US-09-596-746-64	Sequence 64, Appl
20	1526	82.8	395	19	US-09-596-746A-64	Sequence 64, Appl
21	1498	81.3	369	19	US-09-596-746A-40	Sequence 40, Appl
22	1493	81.1	368	19	US-09-596-746-40	Sequence 40, Appl
23	1493	81.1	385	19	US-09-596-746A-56	Sequence 56, Appl
24	1492	81.0	560	19	US-09-596-746-52	Sequence 52, Appl
25	1492	81.0	560	19	US-09-596-746A-52	Sequence 52, Appl
26	1490	80.9	392	19	US-09-596-746-72	Sequence 72, Appl
27	1490	80.9	393	19	US-09-596-746A-72	Sequence 72, Appl
28	1488	80.8	384	19	US-09-596-746-56	Sequence 56, Appl
29	1485	80.6	408	19	US-09-596-746-78	Sequence 78, Appl
30	1485	80.6	409	19	US-09-596-746A-78	Sequence 78, Appl
31	1289.5	70.0	370	19	US-09-596-746A-36	Sequence 36, Appl
32	1284.5	69.7	369	19	US-09-596-746A-36	Sequence 36, Appl
33	1281.5	69.6	393	19	US-09-596-746-66	Sequence 66, Appl
34	1281.5	69.6	394	19	US-09-596-746A-66	Sequence 66, Appl
35	1253.5	68.1	368	19	US-09-596-746A-42	Sequence 42, Appl
36	1248.5	67.8	367	19	US-09-596-746-42	Sequence 42, Appl
37	1245.5	67.6	392	19	US-09-596-746A-74	Sequence 74, Appl
38	1217.5	66.1	370	19	US-09-596-746A-32	Sequence 32, Appl
39	1213.5	65.9	378	19	US-09-596-746-54	Sequence 54, Appl
40	1213.5	65.9	378	19	US-09-596-746A-54	Sequence 54, Appl
41	1212.5	65.8	369	19	US-09-596-746-32	Sequence 32, Appl
42	1209.5	65.7	393	19	US-09-596-746-68	Sequence 68, Appl
43	1209.5	65.7	394	19	US-09-596-746A-68	Sequence 68, Appl
44	1205.5	65.4	401	19	US-09-596-746-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1
US-09-596-746A-24
; Sequence 24, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinstost, Gerald
; APPLICANT: Dykhuiszen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRN
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-24

Query Match 100.0%; Score 1842; DB 19; Length 375;
Best Local Similarity 100.0%; Pred. No. 8; 8e-125;
Matches 375; Conservative 0; Mismatches 0; Gaps 0;
QY 1 MACNNSGKDGNTSANSADSEYKGNLREISKIRDSNVLAVKEVALLSSIDEIAKA 60

```

Db 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDTSNAVLAVKEVALLSSIDEIAAKAI 60
QY 61 IGKIHONNGLDTEYNNHNSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCSSEFTFN 120
Db 61 IGKIHONNGLDTEYNNHNSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCSSEFTFN 120
QY 121 KKEKHTDLDGKEGYTDADAKAIIKTNGTKGAEEGLKFESEVEVLSAAKEMLANSVKE 180
Db 121 KKEKHTDLDGKEGYTDADAKAIIKTNGTKGAEEGLKFESEVEVLSAAKEMLANSVKE 180
QY 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 240
Db 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 240
QY 241 LTSIDELAKAIGKIRKNDVSLDNEADHNSLISGAYLSNLTTRKISAKDSGELKAEIE 300
Db 241 LTSIDELAKAIGKIRKNDVSLDNEADHNSLISGAYLSNLTTRKISAKDSGELKAEIE 300
QY 301 KAKCSEEFPAKLKGEHTDLDGKEGYTDNNAKAIILKTNNDKTKGADELEKLFESVKNLSK 360
Db 301 KAKCSEEFPAKLKGEHTDLDGKEGYTDNNAKAIILKTNNDKTKGADELEKLFESVKNLSK 360
QY 361 AKEMLTNSVKELTS 375
Db 361 AKEMLTNSVKELTS 375

RESULT 2
US-09-596-746-24
; Sequence 24, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-24

```

```

Query Match 99.7%; Score 1837; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 2e-124;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTIDTSNAVLAVKEVALLSSIDEIAAKAI 61
Db 1 ACNNSGKDGNTSANSADSVKGPMLTEISKRTIDTSNAVLAVKEVALLSSIDEIAAKAI 61
QY 62 GKRIHONNGLDTEYNNHNSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCSSEFTFN 120
Db 62 GKRIHONNGLDTEYNNHNSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCSSEFTFN 120
QY 122 LKEKHTDLDGKEGYTDADAKAIIKTNGTKGAEEGLKFESEVEVLSAAKEMLANSVKE 181
Db 122 LKEKHTDLDGKEGYTDADAKAIIKTNGTKGAEEGLKFESEVEVLSAAKEMLANSVKE 181
QY 182 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 241
Db 182 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 241

```

```

Db 181 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 240
QY 242 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAYLSNLTTRKISAKDSGELKAEIEK 301
Db 242 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAYLSNLTTRKISAKDSGELKAEIEK 301
QY 302 AKKCSSEFPALKGEHTDLDGKEGYTDNNAKAIILKTNNDKTKGADELEKLFESVKNLSKA 361
Db 302 AKKCSSEFPALKGEHTDLDGKEGYTDNNAKAIILKTNNDKTKGADELEKLFESVKNLSKA 361
QY 362 AKEMLTNSVKELTS 375
Db 362 AKEMLTNSVKELTS 375

RESULT 3
US-09-596-746-60
; Sequence 60, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-60

```

```

Query Match 99.6%; Score 1834; DB 19; Length 398;
Best Local Similarity 99.7%; Pred. No. 3.6e-124;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTIDTSNAVLAVKEVALLSSIDEIAAKAI 61
Db 25 SCNNSGKDGNTSANSADSVKGPMLTEISKRTIDTSNAVLAVKEVALLSSIDEIAAKAI 84
QY 62 GKRIHONNGLDTEYNNHNSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCSSEFTFN 121
Db 85 GKRIHONNGLDTEYNNHNSLAGAYASTLIKOKLDGKNEGLKEKIDAKKCSSEFTFN 144
QY 122 LKEKHTDLDGKEGYTDADAKAIIKTNGTKGAEEGLKFESEVEVLSAAKEMLANSVKE 181
Db 145 LKEKHTDLDGKEGYTDADAKAIIKTNGTKGAEEGLKFESEVEVLSAAKEMLANSVKE 204
QY 182 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 241
Db 205 LTSPVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTIESNAVLAKEVEETL 264
QY 242 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAYLSNLTTRKISAKDSGELKAEIEK 301
Db 265 TSIDELAKAIGKIRKNDVSLDNEADHNSLISGAYLSNLTTRKISAKDSGELKAEIEK 324
QY 302 AKKCSSEFPALKGEHTDLDGKEGYTDNNAKAIILKTNNDKTKGADELEKLFESVKNLSKA 361
Db 325 AKKCSSEFPALKGEHTDLDGKEGYTDNNAKAIILKTNNDKTKGADELEKLFESVKNLSKA 384
QY 362 AKEMLTNSVKELTS 375
Db 385 AKEMLTNSVKELTS 398

```

```
RESULT 4
US-09-596-746a-60
; Sequence 60, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIORITY FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 399
; TYPE: PRF
; ORGANISM: ospc Chimera
; 09-596-746a-60

Query Match          99.6%; Score 1834; DB 19; Length 399;
Best Local Similarity 99.7%; Pred. No. 3,6e-124;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKTTSDNAVLAVKEVALLSSIDEIAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKTTSDNAVLAVKEVALLSSIDEIAKAI 85
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLKQKLDGLKNEGKKEKIDAKKCEFTFNK 121
DB 86 GKRIHONNGLDTEYNNHNSLAGAVAI STLKQKLDGLKNEGKKEKIDAKKCEFTFNK 145
QY 122 LKEKHTDLGKEGVTDDAKAEILKTNGTKGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 146 LKEKHTDLGKEGVTDDAKAEILKTNGTKGAEELGKLFESVEVLSKAKEMLANSVKE 205
QY 182 LTPVVAESPANVNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVETLL 241
DB 206 LTPVVAESPANVNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVETLL 265
QY 242 TSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISMLITKKSIAIKDSGELKAEIEK 301
DB 266 TSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISMLITKKSIAIKDSGELKAEIEK 325
QY 302 AKKCESEFTAKLGEHTDLGKEGVTDDNAKKAILKTNDKTKGADELEKLFESVKNLSKA 361
DB 326 AKKCESEFTAKLGEHTDLGKEGVTDDNAKKAILKTNDKTKGADELEKLFESVKNLSKA 385
QY 362 AKEMLTNSVKELTS 375
DB 386 AKEMLTNSVKELTS 399

RESULT 5
US-09-596-746-80
; Sequence 80, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
```

```
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRF
; ORGANISM: ospc Chimera
; US-09-596-746-80

Query Match          98.8%; Score 1819.5; DB 19; Length 401;
Best Local Similarity 98.7%; Pred. No. 4,1e-123;
Matches 372; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKTTSDNAVLAVKEVALLSSIDEIAKAI 61
DB 25 SCNNSGKDGNTSANSADSVKGPNTLEISKTTSDNAVLAVKEVALLSSIDEIAKAI 84
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLKQKLDGLKNEGKKEKIDAKKCEFTFNK 121
DB 85 GKRIHONNGLDTEYNNHNSLAGAVAI STLKQKLDGLKNEGKKEKIDAKKCEFTFNK 144
QY 122 LKEKHTDLGKEGVTDDAKAEILKTNGTKGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 145 LKEKHTDLGKEGVTDDAKAEILKTNGTKGAEELGKLFESVEVLSKAKEMLANSVKE 204
QY 182 LTPVVAES--PAMVNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 238
DB 205 LTPVVAESPKRPSVNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 264
QY 239 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISMLITKKSIAIKDSGELKAE 298
DB 265 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISMLITKKSIAIKDSGELKAE 324
QY 299 IEKAKKCESEFTAKLGEHTDLGKEGVTDDNAKKAILKTNDKTKGADELEKLFESVKNL 358
DB 325 IEKAKKCESEFTAKLGEHTDLGKEGVTDDNAKKAILKTNDKTKGADELEKLFESVKNL 384
QY 359 SKAKEMLTNSVKELTS 375
DB 385 SKAKEMLTNSVKELTS 401

RESULT 6
US-09-596-746a-80
; Sequence 80, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 402
; TYPE: PRF
; ORGANISM: ospc Chimera
; US-09-596-746a-80

Query Match          98.8%; Score 1819.5; DB 19; Length 402;
Best Local Similarity 98.7%; Pred. No. 4,1e-123;
Matches 372; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKTTSDNAVLAVKEVALLSSIDEIAKAI 61
```

Db 26 SCNNSGDGTNSANSADESVKGPMLTEISKKTIDSNVLLAVKEVEALLSSIDEIAKAI 85
QY 62 GKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCSSEFTTN 121
Db 86 GKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCSSEFTTN 145
QY 122 LKEKHTDLGEGVYTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 121
Db 146 LKEKHTDLGEGVYTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 205
QY 182 LTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 238
Db 206 LTSPVVAESPKRPSMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 265
QY 239 TLLTSDIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTRKISAKDSGELKA 298
Db 266 TLLTSDIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTRKISAKDSGELKA 325
QY 299 IEKAKKCSSEFTTAKLGEHTDLGEGVYTDNAKKAAILKTNDKTGADELEKLFESVKN 358
Db 326 IEKAKKCSSEFTTAKLGEHTDLGEGVYTDNAKKAAILKTNDKTGADELEKLFESVKN 385
QY 359 SKRAKEMLTNSVKELTS 375
Db 386 SKRAKEMLTNSVKELTS 402

RESULT 7
US-09-596-746-28
; Sequence 28, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Datwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; NO ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-28

Query Match 98.3%; Score 1810.5; DB 19; Length 378;
Best Local Similarity 98.1%; Pred. No. 1.7e-122;
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
QY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVEALLSSIDEIAKA 60
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVEALLSSIDEIAKA 60
QY 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCSSEFTTN 120
Db 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCSSEFTTN 120
QY 121 KLEKHTDLGEGVYTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
Db 121 KLEKHTDLGEGVYTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 237
Db 181 ELTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 237

Db 181 ELTSPVVAESPKRPSMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 240
QY 238 ETLTSDIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTRKISAKDSGELKA 297
Db 241 ETLTSDIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTRKISAKDSGELKA 300
QY 298 IEKAKKCSSEFTTAKLGEHTDLGEGVYTDNAKKAAILKTNDKTGADELEKLFESVKN 357
Db 301 IEKAKKCSSEFTTAKLGEHTDLGEGVYTDNAKKAAILKTNDKTGADELEKLFESVKN 360
QY 358 LSKRAKEMLTNSVKELTS 375
Db 361 LSKRAKEMLTNSVKELTS 378

RESULT 8
US-09-596-746A-28
; Sequence 28, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Datwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-28

Query Match 98.3%; Score 1810.5; DB 19; Length 378;
Best Local Similarity 98.1%; Pred. No. 1.7e-122;
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
QY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVEALLSSIDEIAKA 60
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVEALLSSIDEIAKA 60
QY 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCSSEFTTN 120
Db 61 IGKRIHONGIDTEYNNHNSLAGAYASTLIKOKLDGLKNEGKLEKEIDAAKCSSEFTTN 120
QY 121 KLEKHTDLGEGVYTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
Db 121 KLEKHTDLGEGVYTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVVAES---PAMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 237
Db 181 ELTSPVVAESPKRPSMVNSGKDGNTSANSADSVKGPMLTEISKKTIDSNVLLAVKEVE 240
QY 238 ETLTSDIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTRKISAKDSGELKA 297
Db 241 ETLTSDIDELAKAIGKKIKNDVSLDNEADHNSLISGAYLISNLTTRKISAKDSGELKA 300
QY 298 IEKAKKCSSEFTTAKLGEHTDLGEGVYTDNAKKAAILKTNDKTGADELEKLFESVKN 357
Db 301 IEKAKKCSSEFTTAKLGEHTDLGEGVYTDNAKKAAILKTNDKTGADELEKLFESVKN 360
QY 358 LSKRAKEMLTNSVKELTS 375
Db 361 LSKRAKEMLTNSVKELTS 378

```

RESULT 9
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniaal
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-26

Query Match      86.7%; Score 1597.5; DB 19; Length 374;
Best Local Similarity 87.7%; Pred. No. 4.3e-107;
Matches 328; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVEALLSSIDEIAAKAI 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVEALLSSIDEIAAKAI 60
QY 61 ICKKTHQNNGLDTEYNNHNSLAGAVAI STLTKOKLDGKLNKGLKREKIDAKKCSFTFN 120
DB 61 ICKKTHQNNGLDTEYNNHNSLAGAVAI STLTKOKLDGKLNKGLKREKIDAKKCSFTFN 120
QY 121 KKEKHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVEVLSKAKEMLANSVK 180
DB 121 KKEKHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVE 240
DB 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVE 240
QY 241 LVSIDELA-KAIGKTKINDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAE 299
DB 241 LVSIDELA-KAIGKTKINDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAE 299
QY 300 ERAKKCSSEFTAKLKEGHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVE 359
DB 300 ERAKKCSSEFTAKLKEGHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVE 359
QY 361 KAKKEMLTNSVKEL 373
DB 361 KAKKEMLTNSVKEL 373

RESULT 10
US-09-596-746-26
; Sequence 26, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniaal
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-26

```

```

; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 373
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-26

Query Match      86.5%; Score 1592.5; DB 19; Length 373;
Best Local Similarity 87.7%; Pred. No. 9.9e-107;
Matches 327; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 ACNNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVEALLSSIDEIAAKAI 60
QY 62 GKRIHQNNGLDTEYNNHNSLAGAVAI STLTKOKLDGKLNKGLKREKIDAKKCSFTFN 121
DB 62 GKRIHQNNGLDTEYNNHNSLAGAVAI STLTKOKLDGKLNKGLKREKIDAKKCSFTFN 120
QY 122 LKEKHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVEVLSKAKEMLANSVK 181
DB 122 LKEKHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVE 240
DB 181 ELTSPVAESPAMVNNSGKDGNTSANSADSVKGNLTETISKTTSDSNVAVLAVKEVE 240
QY 242 TSTDELA-KAIGKTKINDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAE 300
DB 242 TSTDELA-KAIGKTKINDVSLDNEADHNSLISGAYLISNLTTRKISAIKDSGELKAE 300
QY 301 KAKKCSSEFTAKLKEGHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVE 360
DB 301 KAKKCSSEFTAKLKEGHTDGLKGEVTDADAKKALTKNGTKTGAEEGLKLFESVE 360
QY 361 KAKKEMLTNSVKEL 373
DB 361 KAKKEMLTNSVKEL 373

RESULT 11
US-09-596-746-62
; Sequence 62, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuizen, Daniaal
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 397
; TYPE: PRP
; ORGANISM: OspC Chimera
US-09-596-746-62

```


Db 385 LAKAKEMLANSVKEL 400

RESULT 14

US-09-596-746A-82
Sequence 82, Application US/09596746A

GENERAL INFORMATION:

APPLICANT: Dettwyler, Raymond J.
APPLICANT: Selinost, Gerald
APPLICANT: Dykhuiszen, Dantial
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 82

LENGTH: 401

TYPE: PR

ORGANISM: ospc Chimera

US-09-596-746A-82

Query Match

Best Local Similarity 85.5%; Score 1575; DB 19; Length 401;
Matches 335; Conservative 22; Mismatches 25; Indels 4; Gaps 2;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKAI 85
QY 62 GKRIHONGLDTEYNNHNSGLAGAVAI STLKQKDLGKNEGKEKIDAAKCSFTFNK 121
DB 86 GKRIHONGLDTEYNNHNSGLAGAVAI STLKQKDLGKNEGKEKIDAAKCSFTFNK 145
QY 122 LKREKHTDLGEGVTDADAKAELIKTNGTKTGABEGLKFESEVLSKAKKEMLANSVKE 181
DB 146 LKREKHTDLGEGVTDADAKAELIKTNGTKTGABEGLKFESEVLSKAKKEMLANSVKE 205
QY 182 LTPVVAES--PAMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 238
DB 206 LTPVVAESPKKPSMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 265
QY 239 TLTLSTIDELA-KAIGKKIKNDVSLDNEADHNSGLISGAVIISLITRKISAIKDSGELKA 297
DB 266 TLTLSTIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAVIISLITRKISAIKDSGELKE 325
QY 298 EIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 357
DB 326 KIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 385
QY 358 LSKRAKEMLANSVKEL 373
DB 386 LAKAKEMLANSVKEL 401

RESULT 15

US-09-596-746-30
Sequence 30, Application US/09596746

GENERAL INFORMATION:

APPLICANT: Dettwyler, Raymond J.
APPLICANT: Selinost, Gerald
APPLICANT: Dykhuiszen, Dantial
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii that Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 377
TYPE: PR
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Ospc Chimera
US-09-596-746-30

Query Match

Best Local Similarity 85.0%; Score 1566; DB 19; Length 377;
Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKA 60
DB 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVEALLSSIDEIAAKA 60
QY 61 IGKRIHONGLDTEYNNHNSGLAGAVAI STLKQKDLGKNEGKEKIDAAKCSFTFNK 120
DB 61 IGKRIHONGLDTEYNNHNSGLAGAVAI STLKQKDLGKNEGKEKIDAAKCSFTFNK 120
QY 121 LKREKHTDLGEGVTDADAKAELIKTNGTKTGABEGLKFESEVLSKAKKEMLANSVK 180
DB 121 LKREKHTDLGEGVTDADAKAELIKTNGTKTGABEGLKFESEVLSKAKKEMLANSVK 180
QY 181 ELTPVVAES--PAMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 237
DB 181 ELTPVVAESPKKPSMVNSGKDGNTSANSADSVKGPNTLEISKRTDSNAVLAVKEVE 240
QY 238 ETLTLSTIDELA-KAIGKKIKNDVSLDNEADHNSGLISGAVIISLITRKISAIKDSGELK 296
DB 241 ETLTLSTIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAVIISLITRKISAIKDSGELK 300
QY 297 AEIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 356
DB 301 EIEKAKKCEEFPAKLGEHTDLGEGVTDNNAKKAILKTNNDKTGADLEKLFESVKN 360
QY 357 NLKRAKEMLANSVKEL 373
DB 361 NLKRAKEMLANSVKEL 377

Search completed: March 18, 2002, 10:08:42
Job time: 970 sec

RESULT 2
US-09-974-992-7

Query Match	31.9%;	Score 587.5;	DB 6;	Length 207;
Best Local Similarity	66.5%;	Pred. No. 2.2e-33;		
Matches 127; Conservative	24;	Mismatches 37;	Indels 3;	Gaps 2;

```

Query Match      9.5%; Score 175; DB 6; Length 1014;
Best Local Similarity 24.3%; Pred. No. 0.00039;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps

QY      8 KDGNTSANSAD-----SVKGPML-----TELSKITDS--NAVLLAVAE 45
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      401 KGTGFADSKGLGVOVAELSTLLEAFQVAKSSLSLEALNATATENEKELTENLNAVTSSEKK 460
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      46 VELLSSIDELIAK-----AIGKIHONNGDITFEYNNHSGLLAGAYASTL--IKQ 94
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      461 LBA-----TVDEYSVKLISESENLLESLRNELNIVQG--KLESTENDLKAAGIQSEVWEKLS 516
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      95 KLDGLKNGKLAKKIDAA-----KNCSEFTYTK-----LKEKH 126
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      517 AEESELOKG--REIDEAATTKRMELALHQSLSDISEHRLQKAMEEFTSNDSEASSLTELK 574
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      127 TDL--CKEGVATADAKAEALLKNTGTFTKCAEELGKTFEVEYVLKRAKEMLANSVKELTSP 185
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      575 RDLBEKITSYEQLAEAGSKSSLEKLEQTLGRTAAEAEVNEKLLKOFDDAOEKSLÖSS 634
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      186 VVAESPANVNSGK-----DGNTSANSADESVKGPMLTE-----ISKRTESNNAVLYAVK 235
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      635 SESELLAFTNNQKIKIOLEELGLIGSGVEKETALKRLEEAIERPNQKETSSEDVLEKIK 694
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      236 EVELLITSIDELA-KAIG-----KKIKN----- 257
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      695 THENQIIEEYKLAHAEAVDTRKVELEDALSKLNESTIEELGAKQOGLKEKESGDLAE 754
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      258 -DVSLDNADHNGSLISGAYLISMLITKKISALDKSGELA-ELEKAKKCSSEPTAKL-- 313
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 755 VNLKLNLELANHGS-----EANELQTKLSALEAKEOTANELSAKTTIEDTLKQLTS 807
QY 314 KGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESYK-NLSKAKEMLTNSVK- 371
Db 808 EGEKLOSOIEKLRVAVAEKSIVLESH-----PEELEKTLSEYKAOQLKEVNAATASVKV 861
QY 372 -ELTS 375
Db 862 AELTS 866

RESULT 5
US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19882
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1018
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1018
; OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19882

Query Match 9.5%; Score 175; DB 6; Length 1018;
Best Local Similarity 24.3%; Pred. No. 0.0004;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;

QY 8 KDGNTSANSAD-----SVKGPNL-----TEISKRTDS-NAVLLAVKE 45
Db 405 KHGETEADSKGYGOVAELQSTLEAFQVYKSSLEALNATENEKELETLNNAVYSEKKK 464
QY 46 VEAALLSIDEIAK-----AIGKKIHONNGLDTEYNNHNSLAGAYASTL--IKO 94
Db 465 LEA---TVDEYSVKISESENLESIRNELNTVG--KLESIENDLKAAGLOESEVMEKKLS 520
QY 95 KIDGLKNEGLKEKIDAA-----KCSERTPTNK-----LKEKH 126
Db 521 AEESELEQKG--REIDEATTKRMELALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTLEKL 578
QY 127 TDL-GKEGYTDADAKAAILKTNGTKGAELGKLFEESEVYLSKAKEMLANSVKELTSP 185
Db 579 RDLBKIKSYEQDLAEASGSSLSKEKLEQTLGRLLAAESVNEKLOEFDQAOEKSLQSS 638
QY 186 VVAESPAMVNNNGK-----DGNTSANSADSEYKGNLT-----ISKKTIESNAVVAVK 235
Db 639 SESELLAETNNQKIKIQLEGLIGSGVEKFTALKRLEAIFERNQKTESSDLYEKIK 698
QY 236 EVELTLTSDIELA-KAIG-----KKIKN----- 257
Db 699 THENOIEEYKRLAHEASGADTRKVELDALSKLMLESTIEELGAKCOGLEKESGDLAE 758
QY 258 -DVSLDNEADHNGSLISGAYILSNLITRKISAIKDSGEILKA-EIEKAKCSEFTAKL-- 313
Db 759 VNLKLNLELANHGS-----EANELQTKLSALEAKEOTANELSAKTTIEDTLKQLTS 811
QY 314 KGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESYK-NLSKAKEMLTNSVK- 371
Db 812 EGEKLOSOIEKLRVAVAEKSIVLESH-----PEELEKTLSEYKAOQLKEVNAATASVKV 865

QY 372 -ELTS 375
Db 866 AELTS 870

RESULT 6
US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881

Query Match 9.5%; Score 175; DB 6; Length 1269;
Best Local Similarity 24.3%; Pred. No. 0.00053;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;

QY 8 KDGNTSANSAD-----SVKGPNL-----TEISKRTDS-NAVLLAVKE 45
Db 656 KHGETEADSKGYGOVAELQSTLEAFQVYKSSLEALNATENEKELETLNNAVYSEKKK 715
QY 46 VEAALLSIDEIAK-----AIGKKIHONNGLDTEYNNHNSLAGAYASTL--IKO 94
Db 716 LEA---TVDEYSVKISESENLESIRNELNTVG--KLESIENDLKAAGLOESEVMEKKLS 771
QY 95 KIDGLKNEGLKEKIDAA-----KCSERTPTNK-----LKEKH 126
Db 772 AEESELEQKG--REIDEATTKRMELALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTLEKL 829
QY 127 TDL-GKEGYTDADAKAAILKTNGTKGAELGKLFEESEVYLSKAKEMLANSVKELTSP 185
Db 830 RDLBKIKSYEQDLAEASGSSLSKEKLEQTLGRLLAAESVNEKLOEFDQAOEKSLQSS 889
QY 186 VVAESPAMVNNNGK-----DGNTSANSADSEYKGNLT-----ISKKTIESNAVVAVK 235
Db 890 SESELLAETNNQKIKIQLEGLIGSGVEKFTALKRLEAIFERNQKTESSDLYEKIK 949
QY 236 EVELTLTSDIELA-KAIG-----KKIKN----- 257
Db 950 THENOIEEYKRLAHEASGADTRKVELDALSKLMLESTIEELGAKCOGLEKESGDLAE 1009
QY 258 -DVSLDNEADHNGSLISGAYILSNLITRKISAIKDSGEILKA-EIEKAKCSEFTAKL-- 313
Db 1010 VNLKLNLELANHGS-----EANELQTKLSALEAKEOTANELSAKTTIEDTLKQLTS 1062
QY 314 KGEHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESYK-NLSKAKEMLTNSVK- 371
Db 1063 EGEKLOSOIEKLRVAVAEKSIVLESH-----PEELEKTLSEYKAOQLKEVNAATASVKV 1116
QY 372 -ELTS 375
Db 1117 AELTS 1121

RESULT 7

```

US-09-708-427-15044
; Sequence 15044, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15044
; LENGTH: 1313
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..1313
; OTHER INFORMATION: Xaa is any amino acid
; OTHER INFORMATION: Ceres Seq. ID 1828627
US-09-708-427-15044

```

```

Query Match
Best Local Similarity 9.2%; Score 169.5; DB 6; Length 1313;
Matches 118; Conservative 60; Mismatches 140; Indels 213; Gaps 23;

```

```

QY 8 KDGNTSANSADSVKGNPLTEISKRTTDSNAV-----LAVKEVEALLSIDEIATAKA 60
DB 6 KTGIMETPRKSPPPRRLSKLSASKSDSNSAPKVRSLRVKGTGLQTLQLOIEDLKA 65
QY 61 ICKIHQNNGLDTEYNNHNSLLAGAVAISTLIKOK-----LDGLK-----NEGIEKI 108
DB 66 -----BEQ-----TELKKAKAIDDLKESELYVEANEKIKEL 101
QY 109 DAARKCSETF-----TNK-----LKEKHT-DLG----- 130
DB 102 AAQKAEESEFEYKFRVLEQAGLEAVOKKDVTSKNELESIRSQHALDISALLTTEEL 161
QY 131 -----KEGYTDADAEAILKTNGTKGAE-----ELGKL-----FES 163
DB 162 QVKEHELSMTADAKKALSHAEATKTAETHAKAEILASLEGRLKALLGSKKEEALEG 221
QY 164 VEVSKAKE-----MLANSVKE-----LTSFVAESPAMVNSG-- 198
DB 222 NEIVSKLSELELGELEKESITLESLEKQGLVBOUKVLEAKAKAES-----CTNSVE 278
QY 199 -----KDGNTSANSADSVKGNPLTEISKRTTESNAV----- 231
DB 279 EMKNKVEHELEKEVEESNRKSSASMSME-----SYMKOLAEHLNHYLHETKSDNAQKEKI 333
QY 232 -LAVKEVEALLSIDE-----LARAIGKIKNDV-----SIDNEADHN 268
DB 334 ELEKTEIEMQRTDLEERYGQVCIACEEASKLENLVESIKSELEISOEKTRALDNEKAT 393
QY 269 GSISGAVYISMLITKIKSAIKDSGELKAELEKAKKSEEFKALGEGHTDGLKEGVTD 328
DB 394 SN-----IQMLDQRTLELIERCKVEBEKSKDMESTLAL-----QENSTES 438
QY 329 NAKKAILKTNNDKTKGADLEKLEFESYKNSKAKE-----MLTNSYKELTS 375
DB 439 SEAKATLLVCOBEIKNC-----ESQVDSLKLASKETNEKYEMKLEDAANEIDS 486

```

```

RESULT 8
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone

```

```

; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7646
; LENGTH: 1881
; TYPE: PRF
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

```

```

Query Match
Best Local Similarity 9.2%; Score 169; DB 7; Length 1881;
Matches 98; Conservative 79; Mismatches 172; Indels 124; Gaps 16;

```

```

QY 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRTTDSNAVLAKEVEALLSIDEIATAKA 60
DB 1411 LAAQLETTSMUSDSTTELEKTELELKVRNLEATSELTSLQNNOSLIELEKTKAA 1470
QY 61 ICKIHQNNGLDTEYNNHNSLLAGAVAISTLIKOKIDGLKNE-----GLKEKIDAAK 114
DB 1471 LTK-----SSKLEVCVGOKSELSDPSLKS-----VKSELKFNENKYNQETTSLKDETEEROK 1523
QY 115 SETFTNKLKEKHTDGLKEGYTDADAEAILKT-----NG 148
DB 1524 IVTLQTELEKDRISEVEKEKAMSENSETVIKESYSDKIKSLESKINSIKENHSEITTHNE 1583
QY 149 TTKGAELEKLEFESYVLSKAEMANSYKELTSPVVAESPAMVNSKDGNTSANS 208
DB 1584 QKTSLKQDIATKLSQHE--SAQTOLEDEKQNKELKASL-----EKHNTESATSI 1631
QY 209 DESVAGPNLTEISKRT-----TESNAVLAKEVEETLLSIDELARAIGK-----K 254
DB 1632 EE--RNNOIKELSETIKSLKTELTSQDALKOSQKERTKTKNSDTESKLEKQEELEK 1669
QY 255 IKNDVSLDNEADHNSLISGAVYISMLITKIKSAIKDSGELK-----ELIEKAK-- 303
DB 1690 VKSPDQRTADE-KLKGITERETALKSELEYKNSGLSTTSELAALKTKYKSLKEKEELOF 1748
QY 304 -----KCSSEFTAKLKEGHTDGLK--EGVTD-----DNAKAILKTNNDKTKGAD 346
DB 1749 LSGKSKLELEDYIOK-----HSDISEKTKALTDLEKERTKQFDOSKSKKLTLELNDLTSTKK 1804
QY 347 ELE-----KLFEVKNLSKAEMLSJNSYKELTS 375
DB 1805 ELEKTEQTSKFKNLEBKDKETIVLKNLELKNLNDNSAKKEISEKVSLES 1857

```

```

RESULT 9
US-10-072-851-15590
; Sequence 15590, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard

```

```

; RESULT 10
; US-09-708-427-15045
; Sequence 15045, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2/50-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15045
; LENGTH: 1304
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

```

```

RESULT 11
US-09-748-875-14
: Sequence 14, Application US/09748875
: GENERAL INFORMATION:
: APPLICANT: BRILES et al.
: TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
: TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
: FILE REFERENCE: 454312-3140
: CURRENT APPLICATION NUMBER: US/09/748,875
: CURRENT FILING DATE: 2000-12-26
: PRIOR APPLICATION NUMBER: 09/298,523
: PRIOR FILING DATE: 1999-04-23
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 14
: LENGTH: 589
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-09-748-875-14

```

Query match 8.6%; Score 158; DB 5; Length 589;

Best Local Similarity 25.1%; Pred. No. 0.0028;
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
OY 19 ESVK-GPMLT-----EISKITDSNAVLAVKEVEALLSSIDEIAAKAIKGIH-QNNG 70
Db 1 EGVSRGNNLTVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 51
OY 71 LDTE-----YNHGSLAGAYAIS-----TLIKOKLDGLKNEGKEKIDA 110
Db 52 LITKLSIEKKKYLVDLKNVLSAEELTSKTEKTEKELATPEOKKOTLPTEPKKVAEA 111
OY 111 AKKSEFTNKLKKEKHD-----LGKE-GVTDADAKKAILKTNGTKTGAEELG 158
Db 112 OKKVEAKKKAEDQEKDRNRPITYKTLELEIAESDVEVKKAELELVYKAKESQDEE 171
OY 159 KLFEES-VEVLSK-----AAKEMLANSVKELTSPVAAS-----190
Db 172 KIKOAEAEVSKOAEATRLKIKITDREAKRKADAKLKEAVEKNVATSEODKPKRRARKG 231
OY 191 -----PAMVNSGKDGNTSANSADSVKGPMLTEISKRTESNMVVLAVKEVELLTS 243
Db 232 VSGELATPDKKENDAKSSDSSV--GEETLPSPLNMANESQTEHR-----KVDDEYI-- 281
OY 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLLISLTKKISAIKDSGELKAEIEKAK 303
Db 282 -----KKMLSEIQLDRR-KHTQNV-----NL-NIKLSAIKTKYLYELSVLKEN 322
OY 304 KCSEFTAKLKGHTDLGKEGVTDNNAKKAAILKTNDKTKGADELKLPESVKNLSKAAK 363
Db 323 SKKEELTSKTKAELT-----AAFEQFKKOTLKEP-----KKVAEAKKVEAKKAKKADOK 372
OY 364 E-----MLTNSVKEL 373
Db 373 EEDRRNYPNTYKTL 387
```

RESULT 12
US-09-748-875-61

; Sequence 61, Application US/09748875
; GENERAL INFORMATION:

; APPLICANT: BRILES et al.

; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

US-09-748-875-61

Query Match 8.6%; Score 158; DB 5; Length 690;
Best Local Similarity 25.1%; Pred. No. 0.0035;

Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
OY 19 ESVK-GPMLT-----EISKITDSNAVLAVKEVEALLSSIDEIAAKAIKGIH-QNNG 70
Db 38 EGVSRGNNLTVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 88
OY 71 LDTE-----YNHGSLAGAYAIS-----TLIKOKLDGLKNEGKEKIDA 110
Db 89 LITKLSIEKKKYLVDLKNVLSAEELTSKTEKTEKELATPEOKKOTLPTEPKKVAEA 148
OY 111 AKKSEFTNKLKKEKHD-----LGKE-GVTDADAKKAILKTNGTKTGAEELG 158
Db 149 OKKVEAKKKAEDQEKDRNRPITYKTLELEIAESDVEVKKAELELVYKAKESQDEE 208
```

OY 159 KLFEES-VEVLSK-----AAKEMLANSVKELTSPVAAS-----190

Db 209 KIKOAEAEVSKOAEATRLKIKITDREAKRKADAKLKEAVEKNVATSEODKPKRRARKG 268

OY 191 -----PAMVNSGKDGNTSANSADSVKGPMLTEISKRTESNMVVLAVKEVELLTS 243

Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPLNMANESQTEHR-----KVDDEYI-- 318

OY 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLLISLTKKISAIKDSGELKAEIEKAK 303

Db 319 -----KKMLSEIQLDRR-KHTQNV-----NL-NIKLSAIKTKYLYELSVLKEN 359

OY 304 KCSEFTAKLKGHTDLGKEGVTDNNAKKAAILKTNDKTKGADELKLPESVKNLSKAAK 363

Db 360 SKKEELTSKTKAELT-----AAFEQFKKOTLKEP-----KKVAEAKKVEAKKAKKADOK 409

OY 364 E-----MLTNSVKEL 373

Db 410 EEDRRNYPNTYKTL 424

RESULT 13
US-09-748-875-1

; Sequence 1, Application US/09748875
; GENERAL INFORMATION:

; APPLICANT: BRILES et al.

; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

US-09-748-875-1

Query Match 8.6%; Score 158; DB 5; Length 691;
Best Local Similarity 25.1%; Pred. No. 0.0035;

Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
OY 19 ESVK-GPMLT-----EISKITDSNAVLAVKEVEALLSSIDEIAAKAIKGIH-QNNG 70
Db 38 EGVSRGNNLTVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 88
OY 71 LDTE-----YNHGSLAGAYAIS-----TLIKOKLDGLKNEGKEKIDA 110
Db 89 LITKLSIEKKKYLVDLKNVLSAEELTSKTEKTEKELATPEOKKOTLPTEPKKVAEA 148
OY 111 AKKSEFTNKLKKEKHD-----LGKE-GVTDADAKKAILKTNGTKTGAEELG 158
Db 149 OKKVEAKKKAEDQEKDRNRPITYKTLELEIAESDVEVKKAELELVYKAKESQDEE 208
OY 159 KLFEES-VEVLSK-----AAKEMLANSVKELTSPVAAS-----190
Db 209 KIKOAEAEVSKOAEATRLKIKITDREAKRKADAKLKEAVEKNVATSEODKPKRRARKG 268
OY 191 -----PAMVNSGKDGNTSANSADSVKGPMLTEISKRTESNMVVLAVKEVELLTS 243
Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPLNMANESQTEHR-----KVDDEYI-- 318
OY 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLLISLTKKISAIKDSGELKAEIEKAK 303
Db 319 -----KKMLSEIQLDRR-KHTQNV-----NL-NIKLSAIKTKYLYELSVLKEN 359
OY 304 KCSEFTAKLKGHTDLGKEGVTDNNAKKAAILKTNDKTKGADELKLPESVKNLSKAAK 363
Db 360 SKKEELTSKTKAELT-----AAFEQFKKOTLKEP-----KKVAEAKKVEAKKAKKADOK 409
```

OY 364 E-----MLNSVKEL 373
 DB 410 EEDRRNYPNTYKTL 424

RESULT 14
 US-09-748-875-62

; Sequence 62, Application US/09748875
 ; GENERAL INFORMATION:
 ; APPLICANT: BRILES et al.
 ; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
 ; FILE REFERENCE: 454312-3140
 ; CURRENT FILING DATE: 2000-12-26
 ; PRIOR APPLICATION NUMBER: 09/298,523
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 62
 ; LENGTH: 701
 ; TYPE: PRP
 ; ORGANISM: Streptococcus pneumoniae
 US-09-748-875-62

Query Match 8.6%; Score 158; DB 5; Length 701;
 Best Local Similarity 25.1%; Pred. No. 0.0035;
 Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

OY 19 ESVK-GPNLT-----EISKITDSNAVLAVEVEALLSSIDEIAKAIGKIH-ONNG 70
 DB 38 EGVRSNNLTVTSSGDISKKYAD-----EVESHLESILKDYKKNLKKYQHTQNG 88
 OY 71 LDTE-----YHNGSLLAGAVAIS-----TLIKOKLDGLKNEGKID 110
 DB 89 LITKLEIKKKYLYDLKVNLSAEELTSKTEKREKLTATFEQFKDTLPTEPEKKVAAE 148
 OY 111 AKKCEFTTNKLEKHTD-----LGKE-GVTDAKAEAILKTNGTKRGAELG 158
 DB 149 OKVEEAKKAEQDQEKDRNYPITYKTLLEIAESDVEKKAELVVKAKESODEE 208
 OY 159 KLPES-VEVLSK-----AAKEMLSYKELTSPVAES-----190
 DB 209 KIKQAEAEVSKQAEATRLKIKITDREAKRKADAKLEAVENKNAVTSSEDDPKRRARG 268
 OY 191 -----PAMVNSGKDGNTSANSADSVKGNLTETSKITESNAVLAVEVEETLTS 243
 DB 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLMANESQTEHR-----KQVDEYI-- 318
 OY 244 IDELAKAIGKIKINDVSLDNEADHNGSLISGAYLISLITKIKISAIKDSGELKAEIEKAK 303
 DB 319 -----KKMLSEIQLDGR-KHTQNV-----NL-NIKLSAIKTKLYELISVLKEN 359
 OY 304 KCSEETAKLIGEHDTLGEKGVTDNNAKKAILKTNNDKTRGADELEKLFESVKNLSKAAK 363
 DB 360 SKKEELTSTKAEILT-----AAFEQFKDTLPKPE---KKVAEAEKKVEAKKAKADOK 409
 OY 364 E-----MLNSVKEL 373
 DB 410 EEDRRNYPNTYKTL 424

RESULT 15
 US-09-748-875-2
 ; Sequence 2, Application US/09748875
 ; GENERAL INFORMATION:
 ; APPLICANT: BRILES et al.
 ; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
 ; FILE REFERENCE: 454312-3140
 ; CURRENT APPLICATION NUMBER: US/09/748,875

; CURRENT FILING DATE: 2000-12-26
 ; PRIOR APPLICATION NUMBER: 09/298,523
 ; PRIOR FILING DATE: 1999-04-23
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 707
 ; TYPE: PRP
 ; ORGANISM: Streptococcus pneumoniae
 US-09-748-875-2

Query Match 8.6%; Score 158; DB 5; Length 707;
 Best Local Similarity 25.1%; Pred. No. 0.0035;
 Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

OY 19 ESVK-GPNLT-----EISKITDSNAVLAVEVEALLSSIDEIAKAIGKIH-ONNG 70
 DB 38 EGVRSNNLTVTSSGDISKKYAD-----EVESHLESILKDYKKNLKKYQHTQNG 88
 OY 71 LDTE-----YHNGSLLAGAVAIS-----TLIKOKLDGLKNEGKID 110
 DB 89 LITKLEIKKKYLYDLKVNLSAEELTSKTEKREKLTATFEQFKDTLPTEPEKKVAAE 148
 OY 111 AKKCEFTTNKLEKHTD-----LGKE-GVTDAKAEAILKTNGTKRGAELG 158
 DB 149 OKVEEAKKAEQDQEKDRNYPITYKTLLEIAESDVEKKAELVVKAKESODEE 208
 OY 159 KLPES-VEVLSK-----AAKEMLSYKELTSPVAES-----190
 DB 209 KIKQAEAEVSKQAEATRLKIKITDREAKRKADAKLEAVENKNAVTSSEDDPKRRARG 268
 OY 191 -----PAMVNSGKDGNTSANSADSVKGNLTETSKITESNAVLAVEVEETLTS 243
 DB 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLMANESQTEHR-----KQVDEYI-- 318
 OY 244 IDELAKAIGKIKINDVSLDNEADHNGSLISGAYLISLITKIKISAIKDSGELKAEIEKAK 303
 DB 319 -----KKMLSEIQLDGR-KHTQNV-----NL-NIKLSAIKTKLYELISVLKEN 359
 OY 304 KCSEETAKLIGEHDTLGEKGVTDNNAKKAILKTNNDKTRGADELEKLFESVKNLSKAAK 363
 DB 360 SKKEELTSTKAEILT-----AAFEQFKDTLPKPE---KKVAEAEKKVEAKKAKADOK 409
 OY 364 E-----MLNSVKEL 373
 DB 410 EEDRRNYPNTYKTL 424

Search completed: March 18, 2002, 09:58:16
 Job time: 374 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:37 ; Search time 68.77 Seconds
(without alignments)
415.377 Million cell updates/sec

Title: US-09-596-746A-24

Perfect score: 1842
Sequence: 1 MACNNSGKDDNTSANSADSE.....KNLSKAKEMLTNSVKELTS 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR.68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	928	50.4	210	2	G70218
2	921	50.0	210	2	S69927
3	879	47.7	210	2	I40144
4	853	46.3	177	2	I40129
5	757.5	41.1	211	2	I40145
6	756.5	41.1	193	2	S70279
7	696.5	37.8	193	2	S70287
8	683.5	37.1	211	2	I40277
9	677.5	36.8	211	2	S69918
10	676.5	36.7	209	2	I40273
11	674.5	36.6	191	2	S70278
12	674.5	36.6	191	2	I40153
13	673.5	36.6	211	2	I40268
14	672.5	36.5	193	2	S70280
15	666	36.2	194	2	S70277
16	665.5	36.1	209	2	S69917
17	665	36.1	190	2	S70273
18	661	35.9	212	2	S70254
19	657.5	35.7	193	2	S70276
20	655.5	35.6	193	2	S70274
21	654.5	35.5	211	2	I40278
22	654	35.5	212	2	S69922
23	653.5	35.5	191	2	S70284
24	652.5	35.4	193	2	S70265
25	652.5	35.4	211	2	S69932
26	652	35.4	194	2	S70268
27	646	35.1	212	2	I40279
28	642.5	34.9	209	2	I40142
29	635	34.5	212	2	I40143

30	634.5	34.4	203	2	I40108	outer surface prot
31	632	34.3	214	2	S69916	outer surface prot
32	631.5	34.3	193	2	S70286	outer surface prot
33	629.5	34.2	191	2	S70288	outer surface prot
34	626	34.0	194	2	S70289	outer surface prot
35	624.5	33.9	179	2	S54222	outer surface prot
36	623.5	33.8	209	2	I40285	outer surface prot
37	621.5	33.7	209	2	I40281	outer surface prot
38	620.5	33.7	178	2	I40125	outer surface prot
39	619	33.6	210	2	S69925	outer surface prot
40	618	33.6	192	2	S70285	outer surface prot
41	617.5	33.5	178	2	I40104	outer surface prot
42	615	33.4	210	2	S69923	outer surface prot
43	614	33.3	212	2	S20543	outer surface prot
44	613	33.3	210	2	S69920	outer surface prot
45	610	33.1	210	2	I40272	outer surface prot

ALIGNMENTS

RESULT 1
G70218
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence: revision 13-Feb-1998 #extl: change 21-Jul-2000
C:Accession: G70218; I40269; S37726; S70281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: G70218
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <KLE>
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AA06329.1; PID:g2689901; TIGR:BB
A:Experimental source: strain B31
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: I40269; MUID:96025162
A:Accession: I40269
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:DA9497; NID:g707092; PIDN:BA008457.1; PID:g769684
R:Jauris-Helpe, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os
A:Reference number: S37726; MUID:93268136
A:Accession: S37726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <TAU>
A:Cross-references: EMBL:x69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dornier, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; MUID:9626448
A:Accession: S70281
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210
A:Cross-references: EMBL:I42887; NID:g858715; PIDN:AA03695.1; PID:g1695212
A:Experimental source: strain Ip2
C:Genetics:
A:Gene: ospC
C:superfamily: Lyme disease spirochete surface protein C

Query Match 50.4%; Score 928; DB 2; Length 210;
Best Local Similarity 98.4%; Pred. No. 1.5e-36;
Matches 186; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60
QY 61 ICKKTHONNGDTEYNHNSLAGAVASTLIKOKLDGKLNKGLKEKIDAAKCKSEFTFN 76
DB 77 ICKKTHONNGDTEYNHNSLAGAVASTLIKOKLDGKLNKGLKEKIDAAKCKSEFTFN 120
QY 121 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180
DB 137 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180
QY 181 ELTSPVVAESP 191
DB 197 ELTSPVVAESP 207

RESULT 2

outer surface protein C precursor - Lyme disease spirochete (strain PKa)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C:Accession: S69927; S72669
J:Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia burgdorferi.
A:Reference number: 140047; MID:95393018
A:Accession: S69927
A:Status: nucleic acid sequence not shown
A:Residues: 1-210 <JAU>
A:Residue type: DNA
A:Cross-references: EMBL:X69589
A:Experimental source: strain PKa
R:Jauris, S.
submitted to the EMBL Data Library, February 1994
A:Reference number: S72669
A:Accession: S72669
A:Molecule type: DNA
A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAW>
A:Cross-references: EMBL:X69589
C:Genetics: none: ospC
Superfamily: Lyme disease spirochete surface protein C

Query Match 50.0%; Score 921; DB 2; Length 210;
Best Local Similarity 97.4%; Pred. No. 3.2e-36;
Matches 186; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 60
QY 61 ICKKTHONNGDTEYNHNSLAGAVASTLIKOKLDGKLNKGLKEKIDAAKCKSEFTFN 76
DB 77 ICKKTHONNGDTEYNHNSLAGAVASTLIKOKLDGKLNKGLKEKIDAAKCKSEFTFN 120
QY 121 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180
DB 137 KLEKHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVK 180
QY 181 ELTSPVVAESP 191
DB 197 ELTSPVVAESP 207

RESULT 3

140144

outer surface protein C precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: 140144; S70282
R:Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A:Title: Expression and sequence of outer surface protein C among North American isolates of Borrelia burgdorferi.
A:Reference number: 140143; MID:95154673
A:Accession: 140144
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: EMBL:U04281; MID:9434663; PIDN:AAQ43297.1; PID:9434664
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease spirochete.
A:Reference number: S70255; MID:96296448
A:Accession: S70282
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210 <LIV>
A:Cross-references: EMBL:L42893; MID:9858721; PIDN:AAQ37001.1; PID:91695218
A:Experimental source: strain 297
C:Genetics: none: ospC
Superfamily: Lyme disease spirochete surface protein C

Query Match 47.7%; Score 879; DB 2; Length 210;
Best Local Similarity 99.4%; Pred. No. 2.8e-34;
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 195 NNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 79
DB 20 NNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVALLSSIDEIAAKA 79
QY 255 IKNDVSLDNEADHNSLSISGAYLSNLTTRKISAIKSGELKAEIKAKKCKSEFTAKLK 314
DB 80 IKNDVSLDNEADHNSLSISGAYLSNLTTRKISAIKSGELKAEIKAKKCKSEFTAKLK 314
QY 315 GEHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVKELT 374
DB 140 GEHTDGLKEGYTDADAKAELIKTKGTGAEELGKLFESVVLAKAKEMLANSVKELT 374
QY 375 S 375
DB 200 S 200

RESULT 4

140129

outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: 140129; S54199
R:Theisen, M.; Borre, M.; Mahliessen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: 140104; MID:95286481
A:Accession: 140129
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-177 <RES>
A:Cross-references: EMBL:X84783; MID:9793825; PIDN:CAA59254.1; PID:9793826
C:Genetics: none: ospC
A:Gene: ospC
Superfamily: Lyme disease spirochete surface protein C

Query Match 46.3%; Score 853; DB 2; Length 177;
Best Local Similarity 98.9%; Pred. No. 3.6e-33;
Matches 175; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 NTSANSADEVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAKAIGKIHONNG 70
 |||||||
 DB 1 NTSANSADEVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAKAIGKIHONNG 60
 QY 71 LDTENVHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFTNKLKEKHDLG 130
 |||||
 DB 61 LDTENVHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFTNKLKEKHDLG 120
 QY 131 KEQVTDADAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNVKELTSPV 187
 |||||||
 DB 121 KRQVTDADAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNVKELTSPV 177

RESULT 5
 140145
 outer surface protein C precursor - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
 C:Accession: I40145
 R:Stevenson, B.; Barthold, S.W.
 Mol. Microbiol. Lett. 124, 367-372, 1994
 A:Title: Expression and sequence of outer surface protein C among North American isolates
 A:Reference number: I40143; MUID:95154673
 A:Accession: I40145
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-211 <RES>
 A:Cross-References: EMBL:U04282; NID:g2314881; PIDN:AA045540.1; PID:g434666
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.1%; Score 757.5; DB 2; Length 211;
 Best Local Similarity 81.2%; Pred. No. 1.1e-28;
 Matches 156; Conservative 13; Mismatches 22; Indels 1; Gaps 1;

QY 1 MACNNGKCGN-TSANSADSVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAK 59
 |||||||
 DB 17 ISCNNGKCGNMASTNPADSVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAK 76
 QY 60 AIGKTIHONNGLDTEYNHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFT 119
 |||||||
 DB 77 AIGKTIHONNGLDTEYNHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFT 136
 QY 120 NKLKEKHTDLGKGVTDADAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNV 179
 |||||
 DB 137 NKLKSHTELGKQDQDDAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNV 196

180 KELTSPVVAESP 191
 |||||||
 197 KELTSPVVAESP 208

RESULT 6
 S70279
 outer surface protein C - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70279
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70279
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <LIV>
 A:Cross-References: EMBL:L42898; NID:g858729; PIDN:AA037007.1; PID:g1695223
 A:Experimental source: strain 25015
 C:Genetics:
 R:Fukunaga, M.; Hamase, A.
 J. Clin. Microbiol. 33, 2415-2420, 1995
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.1%; Score 756.5; DB 2; Length 193;
 Best Local Similarity 82.1%; Pred. No. 1.1e-28;
 Matches 156; Conservative 10; Mismatches 23; Indels 1; Gaps 1;

QY 3 CNNSGKCGN-TSANSADSVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAKAI 61
 |||||||
 DB 1 CNNSGKCGNMASTNPADSVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAKAI 60
 QY 62 GKRIHONNGLDTEYNHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFTNK 121
 |||||||
 DB 61 GKRIHONNGLDTEYNHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFTNK 120
 QY 122 LKEKHNDLGKGVTDADAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNVKE 181
 |||||
 DB 121 LKSHTELGKQDQDDAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNVKE 180

QY 182 LTSPVVAESP 191
 |||||||
 DB 181 LTSPVVAESP 190

RESULT 7
 S70287
 outer surface protein C - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70287
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70287
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <LIV>
 A:Cross-References: EMBL:L42895; NID:g858723; PIDN:AA037003.1; PID:g1695220
 A:Experimental source: strain 28354
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.8%; Score 696.5; DB 2; Length 193;
 Best Local Similarity 76.3%; Pred. No. 6.4e-26;
 Matches 145; Conservative 16; Mismatches 28; Indels 1; Gaps 1;

QY 3 CNNSGKCGN-TSANSADSVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAKAI 62
 |||||||
 DB 1 CNNSGKCGNMASTNPADSVKGPNTLTSKRTSDSNVAVLAKVEEALLSIDETIAAKAI 60
 QY 63 KRIHONNGLDTEYNHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFTNK 121
 |||||
 DB 61 KRIHONNGLDTEYNHNGSLAGAVAYSTLIKOKLDGLKNEGLKEKIDAAKCSSEFTNK 120
 QY 122 LKEKHNDLGKGVTDADAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNVKE 181
 |||||
 DB 121 LKSHTELGKQDQDDAKKALKTNGTKTGABEELGKFESVEVLSKAAKEMLSNVKE 180

QY 182 LTSPVVAESP 191
 |||||||
 DB 181 LTSPVVAESP 190

RESULT 8
 I40277
 outer surface protein C precursor - Borrelia afzelii
 C:Species: Borrelia afzelii
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
 C:Accession: I40277
 R:Fukunaga, M.; Hamase, A.
 J. Clin. Microbiol. 33, 2415-2420, 1995

A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato
A:Reference number: 140269; MUID:96025162
A:Accession: 140277
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: GB:D49501; NID:9707096; PIDN:BAA08461.1; PID:9769688
C:Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 37.1%; Score 683.5; DB 2; Length 211;
Matches 145; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

OY 1 MACNSGKDGNT-SANSADSVKGNLTFEISKRTDSNAVLAVKEVALSSIDEIAAK 59
Db 17 ISCNNSGKGDSTNPDESAGPMLTFEISKRTDSNAVLAVKEVALSSIDEIAAK 59
60 AIGKTIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGLK-EGLEKIDAAKCSFTF 118
Db 77 AIGQKI-QNNGLAADKNGSLAGAYASTLIKOKLDGLKGLGELNKEIAAKKCSAF 135
OY 119 TNKLEKHTDGLKEGYTDADAKEAILKTNGTKTGAEELGLFESVEVLSKAEMTANS 178
Db 136 TKKLDSDNADLKGHNATDADSKAELIKTKNGTKTGAELEELFKSVESLSKAEMTANS 195
OY 179 VKELTSPVAESP 191
Db 196 VKELTSPVAESP 208
|||||

RESULT 9
569918.
Outer surface protein C precursor - Lyme disease spirochete (strain PBr)

C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
A:Variety: strain PBr
C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
R:Journals-Heipre, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, F.
J. Clin. Microbiol. 33, 1860-1866, 1995
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia burgdorferi*
A:Reference number: 140047; MUID:95395018
A:Accession: 569918
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-211 <OAU>
A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 36.8%; Score 677.5; DB 2; Length 211;
Matches 149; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

OY 1 MACNSGKDGNTSANSADSVKGNLTFEISKRTDSNAVLAVKEVALSSIDEIAAK 60
Db 17 ISCNNSGKDGNTSANSADSVKGNLTFEISKRTDSNAVLAVKEVALSSIDEIAAK 60
OY 61 IGGKTIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGLK-NGLEKIDAAKCSFTT 119
Db 76 IGGKTIKNGSLGDEANHNESLAGAYITSLTIKOKLSKNGSGLEKIDAAKCSFTS 135
OY 120 NKLKERTDGLKEGYTDADAKEAILKTNGT-KTGAEELGLFESVEVLSKAEMTANS 178

Db 136 TKLNDHAQLGQGVTDNDNAKKAILKANAGKDGVELEKLSGLESLSKAEMTANS 195
OY 179 VKELTSPVAESP 191
Db 196 VKELTSPVAESP 208
|||||

RESULT 10

140273
Outer surface protein C precursor - *Borrelia afzelii*
C:Species: *Borrelia afzelii*
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C:Accession: 140273
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato
A:Reference number: 140269; MUID:96025162
A:Accession: 140273
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-209 <RES>
A:Cross-references: GB:D49379; NID:91041107; PIDN:BAA08377.1; PID:91041108
C:Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 36.7%; Score 676.5; DB 2; Length 209;
Matches 144; Conservative 20; Mismatches 25; Indels 3; Gaps 3;

OY 1 MACNSGKDGNTSANSADSVKGNLTFEISKRTDSNAVLAVKEVALSSIDEIAAK 60
Db 17 ISCNNSGKGDISTNDDESAGPMLTFEISKRTDSNAVLAVKEVALSSIDEIAAK 75
OY 61 IGGKTIHONNGIDTEYNHNSGLAGAYASTLIKOKLDGLK-EGLEKIDAAKCSFTT 119
Db 76 IGGKI-BANGIGNEDADKNGSLAGAYASTLIKOKLDGLKGLGELNKEIAAKKCSAF 134
OY 120 NKLKERTDGLKEGYTDADAKEAILKTNGTKTGAEELGLFESVEVLSKAEMTANS 179
Db 135 KKLSDSNADLKGHNATDADSKAELIKTKNGTKTGAELEELFKSVESLSKAEMTANS 194
OY 180 KELTSPVAESP 191
Db 195 KELTSPVAESP 206
|||||

RESULT 11

S70278
Outer surface protein C - Lyme disease spirochete
C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: S70278
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in *OspC* variation in Lyme disease spirochete
A:Reference number: S70255; MUID:96296448
A:Accession: S70278
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-191 <LTV>
A:Cross-references: EMBL:L42871; NID:9858738; PIDN:AAB37014.1; PID:91695229
C:Genetics:
A:Experimental source: strain VS461
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 36.6%; Score 674.5; DB 2; Length 191;
Matches 144; Conservative 18; Mismatches 25; Indels 3; Gaps 3;

OY 3 CNNSGKDGNTSANSADSVKGNLTFEISKRTDSNAVLAVKEVALSSIDEIAAKAIG 62

```

||||| : : : ||| ||||| : : : ||||| : : : |||
Db      1 C N N S G K G D G T S A N S A D E S V K G P N L T E I S K R I T D S N A V L A V K E Y E A L L S I D E L - A K T I G 59
Qy      63 K K I H O N N G L D T E Y N H N G S L L A G A V A I S T L I K O K L D G L K N - E G L K E K I D A A K K C S E F T N K 121
      60 K K I - E A N G L O N E A D K N S L L A G A V A I S T L I K O K L D G L K E G L N K E L A E A K K C S E A F T T K 118
Qy      122 L K E K H T D L G K E G V T D A D A K E A I L K T N G T K T K G A E E L G K L F E S V E V L S K A A K E M L A N S V K E 181
      119 L O O S M A D L G K H N A T D A D S K E A I L K T N G T K T K G A E L E F K S V E S L S K A A K E A L S N S V K E 178
Qy      182 L T S P V A E S P 191
      179 L T S P V A E S P 188

RESULT 12
outer surface protein C - Lyme disease spirochete (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
Accession: I40153
Ref: B.P.: McHugh, G.L.; Leong, J.M.; Steere, A.C.
Infect. Immun. 62, 3213-3221, 1994
A:Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in
A:Reference number: I40153; MUID:94314437
A:Accession: I40153
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <RES>
A:Cross-references: EMBL:U08284; NID:g469561; PIDN:AAA21460.1; PID:g469562
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```

```

Query Match      36.6%; Score 674.5; DB 2; Length 191;
Best local Similarity 75.5%; Pred. No. 6.5e-25;
Matches 142; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

Qy      5 N S K R G D G T S A N S A D E S V K G P N L T E I S K R I T D S N A V L A V K E Y E A L L S I D E I A K A I G K K 64
      1 N S K G K G N T S A N S A D E S V K G P N L T E I S K R I T S N A V L A V K E I F T L A S I D E L A T K A I G K K 60
Db      65 I H O N N G L D T E Y N H N G S L L A G A V A I S T L I K O K L D G L K N - E G L K E K I D A A K K C S E F T T K K 123
      61 I O Q N G L A E A G H N G T L L A G A V A I S T L I K O K L D G L K N S E K L E K I E N A K K C S E D F T K K L E 120
Qy      124 E K H T D L G K E G V T D A D A K E A I L K T N G T K T K G A E E L G K L F E S V E V L S K A A K E M L A N S V K E L T 183
      121 G E H A O G I E N V T D E N K K A I L I T D A K D K G A E L E K L F K A V E N L A R A K A K E M L A N S V K E L T 180
Qy      184 S P V A E S P 191
      181 S P V A E S P 188

RESULT 13
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
Accession: I40268
Ref: Margolis, N.; Hogan, D.; Cleplak, W.
Gene 143, 105-110, 1994
A:Title: Homology between Borrelia burgdorferi OspC and members of the family of Borrelia
A:Reference number: I40268; MUID:94259285
A:Accession: I40268
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: GB:L25413; NID:g495735; PIDN:AAA22956.1; PID:g495736
C:Genetics:

```

```

A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match      36.6%; Score 673.5; DB 2; Length 211;
Best local Similarity 74.5%; Pred. No. 8.1e-25;
Matches 143; Conservative 16; Mismatches 32; Indels 1; Gaps 1;

Qy      1 M A C N N S G K D G T S A N S A D E S V K G P N L T E I S K R I T D S N A V L A V K E Y E A L L S I D E I A K A 60
      17 I S C N N S K D G T S A N S A D E S V K G P N L T E I S K R I T D S N A V L A V K E Y E A L L S I D E L A K A 76
Db      61 I G K K I H O N N G L D T E Y N H N G S L L A G A V A I S T L I K O K L D G L K N E G - L K E K I D A A K K C S E F T 119
      77 I G K K I D G N N A L G T D N H N G S L L A G A V A I S A L I T E K L S I N D S G E L K A E I E A K K C S E F T 136
Qy      120 N K L E K H T D L G K E G V T D A D A K E A I L K T N G T K T K G A E E L G K L F E S V E V L S K A A K E M L A N S V 179
      137 K K L S D N O A E I G E I N A T D N A K K A I L K T H N A K D K G A E L V K L S E S V A G L L K A A Q A I L A N S V 196
Qy      180 K E L T S P V A E S P 191
      197 K E L T S P V A E S P 208

RESULT 14
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
Accession: S70280
Ref: Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; MUID:96296448
A:Accession: S70280
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-193 <LIV>
A:Cross-references: EMBL:I42868; NID:g858735; PIDN:AMB37011.1; PID:g1695226
A:Experimental source: strain Z57
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```

```

Query Match      36.5%; Score 672.5; DB 2; Length 193;
Best local Similarity 77.5%; Pred. No. 8.1e-25;
Matches 148; Conservative 9; Mismatches 31; Indels 3; Gaps 3;

Qy      3 C N N S G K D G T S A N S A D E S V K G P N L T E I S K R I T D S N A V L A V K E Y E A L L S I D E I A K A I G 62
      1 C N N S G K D G N T S A N S A D E S V K G P N L T E I S K R I T D S N A V L A V K E Y E A L L S I D E L - A K A I G 59
Db      63 K K I H O N N G L D T E Y N H N G S L L A G A V A I S T L I K O K L D G L K - N E G L K E K I D A A K K C S E F T N K 121
      60 K K I K N D S L G D E A N H N S L L A G A V A I S T L I T O K R I S K L N G S E G L K E K I A A K K C S E E S T K 119
Qy      122 L K E K H T D L G K E G V T D A D A K E A I L K T N G T - K T K G A E E L G K L F E S V E V L S K A A K E M L A N S V K 180
      120 L K N H N Q L G I G V T D E A K K A I L K A N A G A D K V G E E L K L S G S L E S L S K A A K E M L A N S V K 179
Qy      181 E L T S P V A E S P 191
      180 E L T S P V A E S P 190

RESULT 15
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
Accession: S70277

```

R.; Lively, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.

MOL. MICROBIOL. 18, 257-269, 1995

A; Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A; Reference number: S70255; MUID:96296448

A;Accession: S70277

A;Status: nucleic acid sequence not shown

A; molecule type: DNA

A;ResIdues: 1-194 <LIV>

A; Cross-references: EMBL:L42873; NID:g858740; PIDN:AAB37016.1; PID:g1695231

A; Experimental source: strain SIMON

C;Genetics:

A;Gene: ospC

C/Superfamily: Lyme disease spirochete surface protein C

Query Match

36.2%; Score 666; DB 2; Length 194;

Best Local Similarity 74.3%; Pred. No. 1.6e-24;

Matches 142; Conservative 13; Mismatches 34; Indels 2; Gaps 2;

3 C N N S G K D G - N T S A N S A D E S V K G P N L T E I S K I T D S N A V L L A K E V E A L L S S I D E I A K A I 61

1 CUNSCVCCDCESTNDADECAKCDNI ETETCKRTNCSNAEVI AVREVEYET VASCTDEI AAKAT 60

QY 62 GKIHQNGLDTEYNNHNSLAGAYAISTLIKOKLDGLKN-EGLEKIDAKKCEFTFN 120

Db 61 GKKIKNDGTLEENAHNGSLAGAYAISNLIKQKLDGLKGLEGLNKEIAEAKNCSEAFK 120

QY 121 KLEKHTDLGEGVTDADAKEAILKNTGTTKGAEELGKLFESVEVLSKAKEMLANSVK 180

Db 121 KLEKHTDLGKENATDEDAKKAIIKTDATKDKGAEELEKLSSESVASLVKAAQEAALNSVK 180

QY	181	ELTSPVAESP	191

Db 181 ELTSPVAESP 191

Search completed: March 18, 2002, 09:56:46
Job time: 369 sec

Job time: 369 sec

XX Claim 43; Page 83-84; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

XX Sequence 378 AA;

Query Match 100.0%; Score 1859; DB 22; Length 378;
 Best Local Similarity 100.0%; Pred. No. 2, 9e-121;
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 MACNNSGKDGNTSANSADSVKGPNTLTKITDSNAVLAVKEVALLSIDETIAAKA 60
1 macnnsykdgnlsansadesvkgpnlteinkitdsnavllavkeveallsideiaaka 60
61 IGGKIHQNNGLDTENNHNHNSLAGAVYSTLIKOKLDGKNEGLKEKIDAAKCSFTFN 120
61 igkkihqngldtenhngslagayalstlikqldgkneglkeidaakkcsftfn 120
DB 61 igkkihqngldtenhngslagayalstlikqldgkneglkeidaakkcsftfn 120
OY 121 KIKKHHDLGKEGVTDAKAEALIKANGTKTGAEELGKLFESVEVLSKAAKEMLANSVK 180
121 kikkhhdlgkegvtadakaealilkangtkgtgaeelgklfesvevlskaakemlansvk 180
DB 121 kikkhhdlgkegvtadakaealilkangtkgtgaeelgklfesvevlskaakemlansvk 180
OY 181 ELTSPVAESPCKPMSVNNNSGKDGNTSANSADSVKGPNTLTKITDSNAVLAVKEV 240
181 eltspvaaespckpmsvnnnsykdgnlsansadesvkgpnlteiskitdsnavllavkev 240
DB 181 eltspvaaespckpmsvnnnsykdgnlsansadesvkgpnlteiskitdsnavllavkev 240
OY 241 ETLTSTIDELAKAIGKKIKNDVSLDNEADHNSLISGAVLYISMLITRKISAKDSGELKA 300
241 etltstidelakaiyggkikndvslidneadhnslisgavlyismlitrkisakdsgeika 300
DB 241 etltstidelakaiyggkikndvslidneadhnslisgavlyismlitrkisakdsgeika 300
OY 301 EIEKAKCSEETAKLKGHTDLGKEGVTDDNAKKAIIKTNDKTKGADELKLFESVKN 360
301 iekakcseetaktlkghtdlgkegvtddnakkaiiktnndktkgadelklfesvkn 360
DB 301 iekakcseetaktlkghtdlgkegvtddnakkaiiktnndktkgadelklfesvkn 360
OY 361 LKAKAKEMLTNSVKELTS 378
361 lkaakemltnsvkelts 378
DB 361 lkaakemltnsvkelts 378

```

RESULT 2

AAB62738
 AAB62738 standard; Protein: 401 AA.

AC AAB62738;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 80.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYV) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.
 DR N-PSDB; AAF29042.

XX Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -

PS Claim 43; Page 150-151; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

XX Sequence 401 AA;

Query Match 98.7%; Score 1834; DB 22; Length 401;
 Best Local Similarity 98.9%; Pred. No. 1, 7e-119;
 Matches 373; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

2 ACNNSGKDGNTSANSADSVKGPNTLTKITDSNAVLAVKEVALLSIDETIAAKA 61
2 acnnsykdgnlsansadesvkgpnlteinkitdsnavllavkeveallsideiaaka 61
DB 25 scnnsykdgnlsansadesvkgpnlteiskitdsnavllavkeveallsideiaaka 84
OY 62 GKKIHQNNGLDTENNHNHNSLAGAVYSTLIKOKLDGKNEGLKEKIDAAKCSFTFN 121
62 gkkihqngldtenhngslagayalstlikqldgkneglkeidaakkcsftfn 121
DB 62 gkkihqngldtenhngslagayalstlikqldgkneglkeidaakkcsftfn 121
OY 122 LKEKHTDLGKGVYDADAKAEALIKANGTKTGAEELGKLFESVEVLSKAAKEMLANSVK 181
122 lkekhtdlgkgvydadakaealilkangtkgtgaeelgklfesvevlskaakemlansvk 181
DB 145 lkekhtdlgkgvydadakaealilkangtkgtgaeelgklfesvevlskaakemlansvk 204
OY 162 LTPSPVAESPCKPMSVNNNSGKDGNTSANSADSVKGPNTLTKITDSNAVLAVKEV 241
162 ltpspvaaespckpmsvnnnsykdgnlsansadesvkgpnlteiskitdsnavllavkev 241
DB 205 ltpspvaaespckpmsvnnnsykdgnlsansadesvkgpnlteiskitdsnavllavkev 264
OY 242 TLTSTIDELAKAIGKKIKNDVSLDNEADHNSLISGAVLYISMLITRKISAKDSGELKA 301
242 tltstidelakaiyggkikndvslidneadhnslisgavlyismlitrkisakdsgeika 301
DB 265 tltstidelakaiyggkikndvslidneadhnslisgavlyismlitrkisakdsgeika 324
OY 302 IEKAKCSEETAKLKGHTDLGKEGVTDDNAKKAIIKTNDKTKGADELKLFESVKN 361
302 iekakcseetaktlkghtdlgkegvtddnakkaiiktnndktkgadelklfesvkn 361
DB 325 iekakcseetaktlkghtdlgkegvtddnakkaiiktnndktkgadelklfesvkn 384
OY 362 SKAAKEMLTNSVKELTS 378
362 skaakemltnsvkelts 378
DB 385 skaakemltnsvkelts 401

```

RESULT 3

AAB62710
 AAB62710 standard; Protein: 374 AA.

AC AAB62710;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 24.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

XX (UNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX
PI Datwyler RJ, Seinoest G, Dykhulzen D, Luft BJ, Gomes-Solecki M;
XX
XX WPI: 2001-050113/06.
DR N-PSDB; AAF29014.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 78-79; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

Sequence 374 AA:

Query Match 97.1%; Score 1805.5; DB 22; Length 374;
Best Local Similarity 98.1%; Pred. No. 1.5e-117;
Matches 370; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 acnnskgdntsansadesvkgpnltetiskkitdsnavllavkeveallssideiaakai 60
QY 62 GKRIHONNGIDTENNHNHNSGLAGAVASTLIKOKLDGLKNEGIEKIDAAKCSFTFNK 121
DB 61 gkrlhngndtelnhngslagayastllkqldglnegikeidaakcsftfnk 120
QY 122 LKEKHTDLGKEGYTDADAKAIIKANGTKTKGADELGKLFESYEVISKAKEMLANSVKE 181
DB 121 lkekhtdlgkegytdadaeailktngtktkgaeelgklfesevlskaakemlansvke 180
QY 182 LTPVVAESPFRKPSMYNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEVE 241
DB 181 ltpvvaes---pamvnnsgkdntsansadesvkgpnltetiskkitdsnavllavkeve 237
QY 242 TLLTSTIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYVLSNLTITKKAISAKDSGELKAE 301
DB 238 tlltsidelakaignkikndvslndeadhngslisgavylsnlittkkaiskdsgelekae 297
QY 302 IEKAKKCSFEPTAKLKGHEHTDLGKEGYTDNNAKKAIIKTNNDTKGADELEKLFESYKVL 361
DB 298 iekakkcseefaktlkgelhtdlgkegytdnnaakailktndtkgadelkelfesvknl 357
QY 362 SKAAKEMLTNSVKELETS 378
DB 358 skaakemltnsvkelts 374

RESULT 4
AAB62728
ID AAB62728 standard; Protein; 398 AA.
XX
XX AAB62728;
AC
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 60.
DE
XX
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
OS
XX
XX Chimeric - Borrelia sp.
PN
XX WO200078966-A1.

XX 28-DEC-2000.
PD
XX
XX 19-JUN-2000; 2000WO-US16915.
PF
XX
XX 18-JUN-1999; 99US-0140042.
PR
XX
XX (UNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seinoest G, Dykhulzen D, Luft BJ, Gomes-Solecki M;
XX
XX WPI: 2001-050113/06.
DR N-PSDB; AAF29032.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 123; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

Sequence 398 AA:

Query Match 97.0%; Score 1802.5; DB 22; Length 398;
Best Local Similarity 97.9%; Pred. No. 2.5e-117;
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 25 acnnskgdntsansadesvkgpnltetiskkitdsnavllavkeveallssideiaakai 84
QY 62 GKRIHONNGIDTENNHNHNSGLAGAVASTLIKOKLDGLKNEGIEKIDAAKCSFTFNK 121
DB 85 gkrlhngndtelnhngslagayastllkqldglnegikeidaakcsftfnk 144
QY 122 LKEKHTDLGKEGYTDADAKAIIKANGTKTKGADELGKLFESYEVISKAKEMLANSVKE 181
DB 121 lkekhtdlgkegytdadaeailktngtktkgaeelgklfesevlskaakemlansvke 180
QY 182 LTPVVAESPFRKPSMYNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEVE 241
DB 181 ltpvvaes---pamvnnsgkdntsansadesvkgpnltetiskkitdsnavllavkeve 237
QY 242 TLLTSTIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYVLSNLTITKKAISAKDSGELKAE 301
DB 242 tlltsidelakaignkikndvslndeadhngslisgavylsnlittkkaiskdsgelekae 301
QY 302 IEKAKKCSFEPTAKLKGHEHTDLGKEGYTDNNAKKAIIKTNNDTKGADELEKLFESYKVL 361
DB 302 iekakkcseefaktlkgelhtdlgkegytdnnaakailktndtkgadelkelfesvknl 361
QY 362 SKAAKEMLTNSVKELETS 378
DB 382 skaakemltnsvkelts 398

RESULT 5
AAB62713
ID AAB62713 standard; Protein; 377 AA.
XX
XX AAB62713;
AC
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 30.
DE
XX

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX WO200078966-A1.
XX 28-DEC-2000.
XX 19-JUN-2000; 2000WO-US16915.
XX 18-JUN-1999; 99US-0140042.
XX (UUNY) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI; 2001-050113/06.
XX N-PSDB; AAF29017.
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX Claim 43; Page 86-87; 160pp; English.
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX Sequence 377 AA;
SQ
Query Match 86.8%; Score 1614.5; DB 22; Length 377;
Best Local Similarity 87.8%; Pred. No. 2.6e-104;
Matches 331; Conservative 20; Mismatches 25; Indels .1; Gaps 1;
QY 1 MACNSGKDGNTSANSADSVKGNLTETINKRTTDSNAVLAVKEVALSSIDEIAAKA 60
DB 1 macnsgkdgntsansadesvkgpnlteinkrttdsnavllavkevalssideiaaka 60
QY 61 IGGKIQNNGLDTENNHNHNSLLAGAVAI STLTKOKLDLKNBGLKEKIDAKKSETFTN 120
DB 61 igkkihqngldtenhnhnslagayaistlikqldglneglkexidaakksetftn 120
QY 121 KLEKHTDLGKGVTDADAKKALIKANGTKTGAEELKLFESVEVLSKAEMLANSVK 180
DB 121 kkehtldlgkgvtdadakealilkanqtktgaeelgklfesyevlskaakemlansvk 180
QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKRITESNAVLAVKEV 240
DB 181 eltspvvaesppkpsmvnnsdgkdgntsansadesvkgpnlteiskritesnavlavkev 240
QY 241 ELLLSIDELA-KAIGKRTKNDVSLDNEADHNSLISGAYLISNLTIRKISAKIDSGEIK 299
DB 241 elllsidela-kaigkrtkndvsl dneadhnslisgaylislntirakisakidsgelk 299
QY 300 AEIEFAKKCESEFTAKLGEHTDLGKGVTDNNAKKA ILKTNDKTKGADELKLFESVK 359
DB 300 aeiefaKKCESEFTAKLGEHTDLGKGVTDNNAKKA ILKTNDKTKGADELKLFESVK 359
QY 360 NLSKAEMLTNSVKEL 376
DB 360 nlskaemltnsvkel 376
QY 361 nlakaakemlansvkel 377
DB 361 nlakaakemlansvkel 377
RESULT 6
AAB62739
ID AAB62739 standard; Protein; 400 AA.
XX

AC AAB62739;
XX 03-APR-2001 (first entry)
DT Borrelia sp chimeric ospC protein SEQ ID NO: 82.
XX
DE Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX WO200078966-A1.
XX 28-DEC-2000.
XX 19-JUN-2000; 2000WO-US16915.
XX 18-JUN-1999; 99US-0140042.
XX (UUNY) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI; 2001-050113/06.
XX N-PSDB; AAF29043.
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX Claim 43; Page 153; 160pp; English.
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX Sequence 400 AA;
SQ
Query Match 85.5%; Score 1589.5; DB 22; Length 400;
Best Local Similarity 86.7%; Pred. No. 1.5e-102;
Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTTDSNAVLAVKEVALSSIDEIAAKAI 61
DB 2 acnnsdgkdgntsansadesvkgpnlteinkrttdsnavllavkevalssideiaakai 61
QY 62 GKTIHONNGLDPTENNHNHNSLLAGAVAI STLTKOKLDLKNBGLKEKIDAKKSETFTN 121
DB 62 gkkihongldptennhnhnslagayaistlikqldglneglkexidaakksetftn 121
QY 144 IGGKIQNNGLDTENNHNHNSLLAGAVAI STLTKOKLDLKNBGLKEKIDAKKSETFTN 144
DB 144 igkkihqngldtenhnhnslagayaistlikqldglneglkexidaakksetftn 144
QY 181 ELLLSIDELA-KAIGKRTKNDVSLDNEADHNSLISGAYLISNLTIRKISAKIDSGEIK 181
DB 181 elllsidela-kaigkrtkndvsl dneadhnslisgaylislntirakisakidsgelk 181
QY 204 IGGKIQNNGLDTENNHNHNSLLAGAVAI STLTKOKLDLKNBGLKEKIDAKKSETFTN 204
DB 204 igkkihqngldtenhnhnslagayaistlikqldglneglkexidaakksetftn 204
QY 241 ELLLSIDELA-KAIGKRTKNDVSLDNEADHNSLISGAYLISNLTIRKISAKIDSGEIK 241
DB 241 elllsidela-kaigkrtkndvsl dneadhnslisgaylislntirakisakidsgelk 241
QY 264 IGGKIQNNGLDTENNHNHNSLLAGAVAI STLTKOKLDLKNBGLKEKIDAKKSETFTN 264
DB 264 igkkihqngldtenhnhnslagayaistlikqldglneglkexidaakksetftn 264
QY 300 AEIEFAKKCESEFTAKLGEHTDLGKGVTDNNAKKA ILKTNDKTKGADELKLFESVK 300
DB 300 aeiefaKKCESEFTAKLGEHTDLGKGVTDNNAKKA ILKTNDKTKGADELKLFESVK 300
QY 324 IGGKIQNNGLDTENNHNHNSLLAGAVAI STLTKOKLDLKNBGLKEKIDAKKSETFTN 324
DB 324 igkkihqngldtenhnhnslagayaistlikqldglneglkexidaakksetftn 324
QY 361 LSKAEMLTNSVKEL 376
DB 361 lskaeMltnsvkel 376
QY 385 lakaakemlansvkel 400
DB 385 lakaakemlansvkel 400

```
RESULT 7
AAB62711
ID AAB62711 standard; Protein: 373 AA.
XX
AC AAB62711;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 26.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PR 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UNYNY ) UNIV NEW YORK STATE RES. FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR MPI: 2001-050113/06.
XX N-PSDB: AAF29015.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 81; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 373 AA;

Query Match 84.0%; Score 1561; DB 22; Length 373;
Best Local Similarity 85.9%; Pred. No. 1.3e-100;
Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSVKGPNTLEINKKITDTSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 acnnskgdgnstansadesvkgpnltelskkitdsnavllavkeveallssideiaakai 60
OY 62 GKRIHONNGLDTENNHNSGLAGAVAI STLTKQKLDGLKNEGLEKIDAKKCSFTFNK 121
DB 61 gkrihngngldteyhngslagayalstllkqldglkneglekidaakcsftfnk 120
OY 122 LKEKHTDGLGEGVTADAKAEILKANGTKTGAELKTFESVEVLSKAKKEMLANSVKE 181
DB 121 lkehtldlgvegvtadadaeailkngtktgaeelgklfesesvlskaakemlansvke 180
OY 182 LTPSVAESPKKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 241
DB 181 ltpsvaesppkpsmvnnskgdgnstansadesvkgpnltelskrttesnavllavkeve 237
OY 242 TULTSIDELA-KAIGKRTKNDVSLDNEADHNSLSIGAVLISLTKKISAIKDSGELKA 300
DB 238 tllaesidelaikaigkrtkndvslgavlsisltkkisaikdsgeleka 297
OY 301 ELEKAKKCEERTAKLKGHTDLGEGVTDDAKKAILTNDKTKGADELKLFESVKN 360
```

```
DB 298 klenakccsedftcklegelhaqigienvtodenakkaillidcaakdgaaeleklfraven 357
OY 361 LSKRAKEMLTNSVKEL 376
DB 358 lakaakemlansvkel 373

RESULT 8
AAB62729
ID AAB62729 standard; Protein: 397 AA.
XX
AC AAB62729;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 62.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PR 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UNYNY ) UNIV NEW YORK STATE RES. FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR MPI: 2001-050113/06.
XX N-PSDB: AAF29033.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 125-126; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 397 AA;

Query Match 83.8%; Score 1558; DB 22; Length 397;
Best Local Similarity 85.6%; Pred. No. 2.3e-100;
Matches 322; Conservative 23; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSVKGPNTLEINKKITDTSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 25 acnnskgdgnstansadesvkgpnltelskkitdsnavllavkeveallssideiaakai 84
OY 62 GKRIHONNGLDTENNHNSGLAGAVAI STLTKQKLDGLKNEGLEKIDAKKCSFTFNK 121
DB 61 gkrihngngldteyhngslagayalstllkqldglkneglekidaakcsftfnk 144
OY 122 LKEKHTDGLGEGVTADAKAEILKANGTKTGAELKTFESVEVLSKAKKEMLANSVKE 181
DB 121 lkehtldlgvegvtadadaeailkngtktgaeelgklfesesvlskaakemlansvke 204
OY 182 LTPSVAESPKKPSMVNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVE 241
DB 181 ltpsvaesppkpsmvnnskgdgnstansadesvkgpnltelskrttesnavllavkeve 261
```

QY 242 TLTSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELKA 300
 Db 262 TLTSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELKA 300
 QY 301 ELERAKKSEEFATKKEGHTDLGEGYTDNNAKKAIIKTNDKTKGADLEKLFESYKN 360
 Db 322 KLENKKSEDFTKKIEGHAGLIGYTDENAKKAILLTDKADKGALEKLFESYKN 360
 QY 361 LSKAKKEMLTNSVKEL 376
 Db 382 LAKAKEMLTNSVKEL 397

RESULT 9
 AAB62715
 ID AAB62715 standard; Protein: 370 AA.

AAB62715:

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 34.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000MO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNY) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M.

PT Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 91-92; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

XX Sequence 370 AA;

Query Match 82.2%; Score 1528.5; DB 22; Length 370;
 Best Local Similarity 86.3%; Pred. No. 2.3e-98;
 Matches 327; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETNKITDSNNAVLLAVEVEALLSSIDEIAAKAI 61
 Db 1 acnnskgdntsansadesvkgpnlteiskitdsnavllaveveallssidel-akai 59

QY 62 GKTHONNGJDTENNHNHSLAGAYISTLIKOKLGLK-NEGICKRIDAKKSEFTFN 120
 Db 60 gkthnngjdtennhnslagayistlitqklkslsgslgkkaakcseefst 119

QY 121 KLEKHTDLGEGYTDADAKAELIKANGT-KTGAEELGKLFESYVLSKAKKEMLTNSV 179

Db 120 KLEKHTDLGEGYTDADAKAELIKANGT-KTGAEELGKLFESYVLSKAKKEMLTNSV 179
 QY 180 KELTSPVAESPCKPSPVNNNSGKDGNTSANSADSVKGNLTETNKITDSNNAVLLAVE 239
 Db 180 keltspvnhg-----msgkdntsansadesvkgpnlteiskitdsnavllave 231
 QY 240 VETLLSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELK 299
 Db 232 VETLLSIDELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLTITTKISAIRKDSGELK 291
 QY 300 AEIERAKKSEEFATKKEGHTDLGEGYTDNNAKKAIIKTNDKTKGADLEKLFESYKN 359
 Db 292 aeieakkkseefatklkeghagligytdenakkailltdkdgaleklfesv 351
 QY 360 NLSKAKKEMLTNSVKELTS 378
 Db 352 nlskaakemltnsvkelt 370

RESULT 10
 AAB62730
 ID AAB62730 standard; Protein: 394 AA.

AAB62730:

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 64.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000MO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNY) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M.

PT Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 128-129; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

XX Sequence 394 AA;

Query Match 82.1%; Score 1525.5; DB 22; Length 394;
 Best Local Similarity 86.0%; Pred. No. 4e-98;
 Matches 326; Conservative 11; Mismatches 31; Indels 11; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETNKITDSNNAVLLAVEVEALLSSIDEIAAKAI 61
 Db 25 scnsgkdntsansadesvkgpnlteiskitdsnavllaveveallssidel-akai 83

```

Oy 62 GKKT HONNGDTENNHNGSSLLAGAVYISTFIKOKLDGK -NEELKEKIDAAKKCSEFTN 120
Db 84 gkktkndgsldneanrnesllagayltstlctqkkslingseglkekiaakkcseeft 143
Oy 121 KLEKHTDLDGKEGVTDAKAEALILKNGT -KTGAELGKLFESYEVLSKAAKEMLANSV 179
Db 144 klkdnhaqjlglygtldenakkaillkanaagkdkgyveeleklsgslsleslskaakemlanv 203
Oy 180 KELTSPVVAESPKKRPSMVNNSGKDGNTSANSADEESYKGRMLTITISKTKTESNAVYLAWE 239
Db 204 keltspvvhg-----nsgkdgnltsanaadesvkgpnlleiskkltesnavylawe 255
Oy 240 VETLTSTIDELAKAIGKKIKNDVSLDNEADHNGSLISGAVLISNLTKTKISATKDSGEK 299
Db 256 veltlstidelakagtkkikndvslndeahngslisgavylisnlttkltsatksgekl 315
Oy 300 AEIEKAKKCEEFYAKLGEHTDLDGKEGVTDDNAKKAILLKTNDKTKGADLEKLFESYK 359
Db 316 aeiekakkcseefaklgehtldlgkevtdnakkailktndckkgadeleklfesvk 375
Db 360 NLSKAAKEMLTNSYKELTS 378
Db 376 nlskaakemltnsykelts 394

RESULT 11
AAB62726
ID AAB62726 standard; Protein; 384 AA.
XX
AC AAB62726;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 56.
XX
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PE 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
XX (UNYNY ) UNIT NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selmost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WP1; 2001-050113/06.
XX
DR N-PSDB; AAF29030.
XX
PT Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS
XX Claim 43: Page 117-118; 160pp: English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SO Sequence 384 AA;

```

[illegible]

	RESULT	12
AA62737	ID	AA62737 standard; Protein; 408 AA.
XX		
AC	AA62737;	
XX		
DT	03-APR-2001	(first entry)
XX		
DE	Borrelia sp chimeric ospc protein SEQ ID NO: 78.	
XX		
KW	Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.	
XX		
OS	Chimeric - Borrelia sp.	
OS	Chimeric - Borrelia sp.	
XX		
PN	WO200078966-A1.	
XX		
PD	28-DEC-2000.	
XX		
PE	19-JUN-2000; 2000WO-US16915.	
XX		
PR	18-JUN-1999; 99US-0140042.	
XX		
PA	(UYNV) UNITV NEW YORK STATE RES FOUND.	
XX	(BROO-) BROOK BIOTECHNOLOGIES INC.	
PI	Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;	
DR	WPI; 2001-050113/06.	
XX	N-PSDB; AAF29041.	
PT		
PT	Compositions of OspC polypeptides from strains of Borrelia which cause	
XX	Lyme disease are used to immunize animals and detect immune responses	
XX	to Lyme disease -	
XX		
XX	Claim 43; Page 147-148; 160pp; English.	
CC	The present invention provides compositions comprising ospc proteins and	
CC	chimeric ospc proteins from members of the Borrelia genus. These may be	
CC	Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as	
CC	vaccines against Borrelia infection, which is spread by ticks and leads	

CC to Lyme disease.
XX
SQ Sequence 408 AA:

Query Match 80.3%; Score 1493.5; DB 22; Length 408;
Best Local Similarity 81.7%; Pred. No. 7e-96;
Matches 308; Conservative 33; Mismatches 33; Indels 3; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEVKGPNLTETIKKTTDSNAVLLAVKEVALLSSIDEIATAKAI 61
DB 25 scmsngkdgntsaansadesvkgpnlteiskitdsnavllavkeveallssidelataka 84
OY 62 GKRIHNNGLDPTENNNGSLAGAVAIPTLIKOKLIDGLKNEGKKEIDAKKSEFTFNK 121
DB 85 gkrihngldpteynhngslagayaistlikgkldglnegkkekidaakkseftfnk 144
OY 122 LKEKHTDLGEGVTDADAKAEILKANGTGTGAELGKLFESVEVLSKAKEMLANSVKE 181
DB 145 lkehtldlgvtdadakeailkngtktgaeelgklfesevevlskaakemlansvke 204
OY 182 LTSPVAESPCKRPMWNNNGKDGNTSANSADSEVKGPNLTETISKTTESNAVYLAKEYE 241
DB 205 ltspvaespckrpfngmsg--gdsastnpdesakypnlvtviskkitdsnaflavkeye 262
OY 242 TLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTETKRISAIKDSGELKAE 301
DB 263 allssidelakelgkrikndvslneadhnngslisgaysliskltlqklsavl-nseelk 321
OY 302 IEKAKKSEEFPAKLKGEHTDLGEGVTDNNAKKAILTNNDKTGADLEKLFESVKUL 361
DB 322 lkeakdcsgkftklkdsnaelqisvgddnakkaillkthgkdkyakeleelkfslesl 381
OY 362 SKAKEMLTNSVKELTIS 378
DB 382 skaagaaltnsvkeltn 398

RESULT 13

AAB62718
ID AAB62718 standard; Protein; 368 AA.

AC AAB62718;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 40.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

XX N-PSDB; AAF29022.

XX Compositions of ospC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PI to Lyme disease -
XX

PS Claim 43; Page 99-100; 160pp; English.

XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

SQ Sequence 368 AA:

Query Match 80.1%; Score 1489.5; DB 22; Length 368;
Best Local Similarity 82.5%; Pred. No. 1.2e-95;
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSEVKGPNLTETIKKTTDSNAVLLAVKEVALLSSIDEIATAKAI 61
DB 1 acnngkdgntsaansadesvkgpnlteiskitdsnavllavkevetllasidelataka 60
OY 62 GKRIHNNGLDPTENNNGSLAGAVAIPTLIKOKLIDGLKNEGKKEIDAKKSEFTFNK 121
DB 61 gkrihngldpteynhngslagayaistlikgkldglnegkkekidaakkseftfnk 119
OY 122 LKEKHTDLGEGVTDADAKAEILKANGTGTGAELGKLFESVEVLSKAKEMLANSVKE 181
DB 120 lkehtldlgvtdadakeailkngtktgaeelgklfesevevlskaakemlansvke 179
OY 182 LTSPVAESPCKRPMWNNNGKDGNTSANSADSEVKGPNLTETISKTTESNAVYLAKEYE 241
DB 180 ltspvthg-----nsgkdgntsaansadesvkgpnlteiskitdsnavllavkeye 231
OY 242 TLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTETKRISAIKDSGELKAE 301
DB 232 tltssidelakelgkrikndvslneadhnngslisgaysliskltlqklsavl-nseelk 291
OY 302 IEKAKKSEEFPAKLKGEHTDLGEGVTDNNAKKAILTNNDKTGADLEKLFESVKUL 361
DB 292 lkekacseefpklkgehtdlggevtdnnaakkailltnndktgadelkelfesvkn 351
OY 362 SKAKEMLTNSVKELTIS 378
DB 352 skaakemltnsvkelts 368

RESULT 14

AAB62734
ID AAB62734 standard; Protein; 392 AA.

AC AAB62734;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 72.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

Tue Mar 19 10:57:44 2002

us-09-596-746a-28.rag

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:37 ; Search time 124.19 Seconds
(without alignments)
441.679 Million cell updates/sec

Title: US-09-596-746a-24

Perfect score: 1842
Sequence: 1 MACNNGSGKQNTSANSADAES.....KNLSKAKEMLINSVKELTS 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organella:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	926	50.3	192	2	09S3P3
2	921	50.0	192	2	09R7B1
3	910	49.4	191	2	P70818
4	908	49.3	193	2	09RR53
5	893	48.5	200	2	09ROR9
6	885	48.0	182	2	09R7B2
7	885	48.0	191	2	031120
8	879	47.7	192	2	09S3P2
9	879	47.7	210	2	044719
10	853	46.3	177	2	044999
11	843.5	45.8	178	2	09REH7
12	829	45.0	175	2	09R7B4
13	802	43.5	163	2	09R7A9
14	786	42.7	165	2	09AGB1
15	773	42.0	159	2	09R7B0
16	757.5	41.1	211	2	044720
17	756.5	41.1	193	2	P94237
18	707	38.4	193	2	031115
19	696.5	37.8	193	2	P94234

20	683.5	37.1	211	2	049576
21	678.5	36.8	194	2	031122
22	678.5	36.8	211	2	044977
23	677	36.8	212	2	09KIM5
24	676.5	36.7	209	2	044671
25	675	36.6	201	2	P96571
26	674.5	36.6	191	2	044726
27	674.5	36.6	191	2	P94245
28	674	36.6	201	2	P96573
29	673.5	36.6	211	2	044976
30	672.5	36.5	193	2	P94242
31	672.5	36.5	202	2	P96508
32	672	36.5	203	2	050624
33	672	36.5	212	2	044705
34	670	36.4	201	2	P96572
35	667	36.2	203	2	050619
36	666	36.2	194	2	P94247
37	665	36.1	190	2	P94244
38	661	35.9	212	2	044727
39	660.5	35.9	201	2	09ROR8
40	658.5	35.7	194	2	031114
41	657.5	35.7	193	2	P94231
42	657.5	35.7	211	2	0926C7
43	655.5	35.6	193	2	P94233
44	655.5	35.6	209	2	044883
45	654.5	35.5	211	2	049577

ALIGNMENTS

RESULT	ID	Q9S3P3	PRELIMINARY:	PRT:	192 AA.
AC	Q9S3P3	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)				
DE	OUTER SURFACE PROTEIN C (FRAGMENT).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-IP2;				
RX	MEDLINE-96296448; PubMed-8709845;				
RA	Lively I., Gibbs C.P., Schuster R., Dorner F.;				
RT	"Evidence for lateral transfer and recombination in OspC variation in Lyme disease Borrelia."				
RL	Mol. Microbiol. 18:257-269(1995).				
DR	EMBL: L42887; AAB36995.1;				
DR	InterPro: IPR001800; Lipoprotein_6.				
DR	Pfam: PF01441; Lipoprotein_6; 1.				
FT	NON_TER 1 192				
FT	NON_TER 1 192				
SQ	SEQUENCE 192 AA; 20287 MW; 118467AC84C7E3D CRC64;				
Query Match					
Best Local Similarity 99.3%; Score 926; DB 2; Length 192;					
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	3	CNNNGKDGNTSANSADSVKGNLTKITDSTNVLAVKEVEALLSSIDEIAAKAIG 62			049576 borrelia af
DB	1	CNNNGKDGNTSANSADSVKGNLTKITDSTNVLAVKEVEALLSSIDEIAAKAIG 60			031122 borrelia bu
QY	63	KTIHNNGLDTEYNNHNSLAGAIASTLTKOKLDGKNGLEKEDAKKCEFTTNKL 122			044977 borrelia bu
DB	61	KTIHNNGLDTEYNNHNSLAGAIASTLTKOKLDGKNGLEKEDAKKCEFTTNKL 120			09KIM5 borrelia af
QY	123	KEHTDGLGKEGYVDADAKKAILTKNGTKRGABELCKLFESVVLAKAKEMLANSVKL 182			044671 borrelia af

DB 121 KEKHTDLGEGVTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVKEL 180
 QY 183 TSPVVAESP 191
 DB 181 TSPVVAESP 189

RESULT 2

O9R7B1

ID O9R7B1 PRELIMINARY: PRT: 192 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT)
 OS Borrelia burgdorferi (Lyme disease spirochete).
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;

[1] SEQUENCE FROM N.A.

RA STRAIN-TESTS:

RC MEDLINE-97478003; PubMed-9336916;

RA Ras N.M., Postic D., Foretz M., Baranton G.;

RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the

RU U.S.A.'";

RN Int. J. Syst. Bacteriol. 47:1112-1117(1997).

[2] SEQUENCE FROM N.A.

RA STRAIN-TESTS:

RC Marit-Ras N., Postic D., Foretz M., Baranton G.;

RU Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL; U91798; AAB81895.1; -

DR InterPro: IPR001800; Lipoprotein_6.

DR Pfam: PF01441; Lipoprotein_6; 1.

DR ProDom: PD001149; Lipoprotein_6; 1.

FT NON_TER 1

FT NON_TER 192

SQ SEQUENCE 192 AA; 20297 MW; 6770502A20AAE764 CRC64;

Query Match 50.0%; Score 921; DB 2; Length 192;
 Best Local Similarity 98.4%; Pred. No. 3.5e-35;
 Matches 187; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAKA 60

QY 3 ISCNNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAKA 62

QY 61 ICKKIHNNGLDTEYHNHNSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSFTFTN 120

QY 63 ICKKIHNNGLDTEYHNHNSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSFTFTN 122

QY 121 KKEKHTDLGEGVTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180

QY 123 KKEKHTDLGEGVTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 182

QY 181 ELTSPVVAES 190

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

QY 183 ELTSPVVAES 192

OX NCBI_TaxID=139;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN-2-1498 CA4.
 RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
 RU Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; L81131; AAB06569.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 193
 SQ SEQUENCE 191 AA; 20126 MW; D2B9B1C82B4DC3C0 CRC64;

Query Match 49.4%; Score 910; DB 2; Length 191;
 Best Local Similarity 98.9%; Pred. No. 1.1e-34;
 Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 NNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAKAIGK 63

QY 1 NNSGKDGNTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAKAIGK 60

QY 64 KIHONNGLDTEYHNHNSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSFTFTNKLK 123

QY 61 KIHONNGLDTEYHNHNSLAGAVASTLIKOKLDGLKNEGLKEKIDAAKCSFTFTNKLK 120

QY 124 EKHHTDLGEGVTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVKELT 183

QY 121 EKHHTDLGEGVTDADAKAAILKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVKELT 180

QY 184 SPVVAESP 191

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

QY 181 SPVVAESP 188

```
Db 67 ICKIKQNNGLDTEENHNSLAGAIAISTLIKQKLDGKLNELKKEIDAKKCSFTFN 126
Qy 121 KKEKHTDGLKRECVTDADAKKALIKTNGTKTGAELGKLFESVEVLSRAKAKMLANSVK 180
Db 127 KKEKHTDGLKRECVTDADAKKALIKTNGTKTGAELGKLFESVEVLSRAKAKMLANSVK 186
Qy 181 ELTSPV 187
Db 187 ELTSPV 193

RESULT 5
Q9R09 PRELIMINARY: PRT: 200 AA.
ID Q9R09;
AC Q9R09;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OSPC.
Borrelia burgdorferi (Lyme disease spirochete).
Plasmid cp26.
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID-139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2E7;
RA MEDLINE-20002545; PubMed-10531219;
RA Hofmeister E.R., Glas G.E., Childs J.E., Persing D.H.;
RT "Population dynamics of a naturally occurring heterogeneous mixture of
RL infect. Immun. 67:5709-5716(1999).
DR EMBL: AF074464; AAD23911.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BEB6B CRC64;

Query Match 48.5%; Score 893; DB 2; Length 200;
Best Local Similarity 98.4%; Pred. No. 6.7e-34;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDETIAKA 60
Db 17 ICKIKQNNGLDTEENHNSLAGAIAISTLIKQKLDGKLNELKKEIDAKKCSFTFN 76
Qy 61 ICKIKQNNGLDTEENHNSLAGAIAISTLIKQKLDGKLNELKKEIDAKKCSFTFN 120
Db 77 ICKIKQNNGLDTEENHNSLAGAIAISTLIKQKLDGKLNELKKEIDAKKCSFTFN 136
Qy 121 KKEKHTDGLKRECVTDADAKKALIKTNGTKTGAELGKLFESVEVLSRAKAKMLANSVK 180
Db 137 KKEKHTDGLKRECVTDADAKKALIKTNGTKTGAELGKLFESVEVLSRAKAKMLANSVK 196
Qy 181 ELTSPV 187
Db 187 ELTSPV 200

RESULT 6
Q9R7B2 PRELIMINARY: PRT: 182 AA.
ID Q9R7B2;
AC Q9R7B2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
```

```
OX NCBI_TaxID-139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RX MEDLINE-97478003; PubMed-9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91797; AAB1894.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match 48.0%; Score 885; DB 2; Length 182;
Best Local Similarity 99.5%; Pred. No. 1.4e-33;
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GKDNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDETIAKAIGKKIH 66
Db 1 GKDNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDETIAKAIGKKIH 60
Qy 67 QNNGLDTEYNNHNSLAGAIAISTLIKQKLDGKLNELKKEIDAKKCSFTFNKLEKH 126
Db 61 QNNGLDTEYNNHNSLAGAIAISTLIKQKLDGKLNELKKEIDAKKCSFTFNKLEKH 120
Qy 127 TDGKRECVTDADAKKALIKTNGTKTGAELGKLFESVEVLSRAKAKMLANSVKELTSPV 186
Db 121 TDGKRECVTDADAKKALIKTNGTKTGAELGKLFESVEVLSRAKAKMLANSVKELTSPV 180
Qy 187 VA 188
Db 181 VA 182

RESULT 7
Q031120 PRELIMINARY: PRT: 191 AA.
ID Q031120;
AC Q031120;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID-139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OC10;
RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF029869; AAB8652.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER
SQ SEQUENCE 191 AA; 20502 MW; E4FB56BCB61740F9 CRC64;

Query Match 48.0%; Score 885; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.5e-33;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 195 NNSGKGNTSANSADSVKGNLTETSKITETSNVAVLAVKEVEETLLTSLDELAKAIGKK 254
 DB 10 NNSGKGNTSANSADSVKGNLTETSKITETSNVAVLAVKEVEETLLTSLDELAKAIGKK 69
 QY 255 IKNDVSLDNEADHNSLISGAVLISNLTTRKISAIKDSGELKAEIERAKKCSSEFTAKLK 314
 DB 70 IKNDVSLDNEADHNSLISGAVLISNLTTRKISAIKDSGELKAEIERAKKCSSEFTAKLK 129
 QY 315 GEHTDLGKEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELT 374
 DB 130 GEHTDLGKEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELT 189
 QY 375 S 375
 DB 190 S 190

QY 8
 ID 09S3P2 PRELIMINARY; PRT; 192 AA.
 AC 09S3P2;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RX MEDLINE=96296448; PubMed=8709845;
 RA Livey I., Gibbs C.P., Schuster R., Dornier F.;
 RT "Evidence for lateral transfer and recombination in OSCP variation in
 RT Lyme disease Borrelia."
 RL Mol. Microbiol. 18:257-269(1995).
 DR EMBL: LA2893; AAB37001.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT 192
 SQ SEQUENCE 192 AA; 20472 MW; 46AC8F93E4DFED6C CRC64;

Query Match 47.7%; Score 879; DB 2; Length 192;
 Best Local Similarity 99.4%; Pred. No. 2,8e-33;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 195 NNSGKGNTSANSADSVKGNLTETSKITETSNVAVLAVKEVEETLLTSLDELAKAIGKK 254
 DB 2 NNSGKGNTSANSADSVKGNLTETSKITETSNVAVLAVKEVEETLLTSLDELAKAIGKK 61
 QY 255 IKNDVSLDNEADHNSLISGAVLISNLTTRKISAIKDSGELKAEIERAKKCSSEFTAKLK 314
 DB 62 IKNDVSLDNEADHNSLISGAVLISNLTTRKISAIKDSGELKAEIERAKKCSSEFTAKLK 121
 QY 315 GEHTDLGKEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELT 374
 DB 132 GEHTDLGKEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELT 181
 QY 375 S 375
 DB 182 S 182

RESULT 9
 ID 044719 PRELIMINARY; PRT; 210 AA.
 AC 044719;
 DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 GN OSCP.
 DE OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB19;
 RX MEDLINE=95154673; PubMed=7851744;
 RA Stevenson B., Barthold S.W.;
 RT "Expression and sequence of outer surface protein C among North
 RT American isolates of Borrelia burgdorferi."
 RL FEMS Microbiol. Lett. 124:367-372(1994).
 DR EMBL: U04281; AAC43297.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 210 AA; 22526 MW; 2D672991D584E4EE CRC64;

Query Match 47.7%; Score 879; DB 2; Length 210;
 Best Local Similarity 99.4%; Pred. No. 3,1e-33;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 195 NNSGKGNTSANSADSVKGNLTETSKITETSNVAVLAVKEVEETLLTSLDELAKAIGKK 254
 DB 20 NNSGKGNTSANSADSVKGNLTETSKITETSNVAVLAVKEVEETLLTSLDELAKAIGKK 79
 QY 255 IKNDVSLDNEADHNSLISGAVLISNLTTRKISAIKDSGELKAEIERAKKCSSEFTAKLK 314
 DB 80 IKNDVSLDNEADHNSLISGAVLISNLTTRKISAIKDSGELKAEIERAKKCSSEFTAKLK 139
 QY 315 GEHTDLGKEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELT 374
 DB 140 GEHTDLGKEGVTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAKEMLTNSVKELT 199
 QY 375 S 375
 DB 200 S 200

RESULT 10
 ID 044999 PRELIMINARY; PRT; 177 AA.
 AC 044999;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TXGW;
 RX MEDLINE=95286481; PubMed=7768799;
 RA Theisen M., Botre W., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 RA Hansen K.;
 RT "Evolution of the Borrelia burgdorferi outer surface protein OSCP."
 RT J. Bacteriol. 177:3036-3044(1995).
 RL EMBL: X84783; CAA59254.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT 177
 SQ SEQUENCE 177 AA; 18800 MW; 156671B9614E7A2D CRC64;

Query Match 46.3%; Score 853; DB 2; Length 177;

Query Match 43.5%; Score 802; DB 2; Length 163;
 Best Local Similarity 98.8%; Pred. No. 6,9e-30;
 Matches 161; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACNNSGDNRTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAAKAI 61
 DB 1 SCNNSGDNRTSANSADSVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAAKAI 60

QY 62 GKRIHONGDTEYNHNSGLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFTTNK 121
 DB 61 GKRIHONGDTEYNHNSGLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFTTNK 120

QY 122 LKEKHTDLGKEGYVDADAKAELIKTNGTKTGAEELGKLFESV 164
 DB 121 LKEKHTDLGKEGYVDADAKAELIKTNGTKTGAEELGKLFESV 163

RESULT 14
 DBI
 AC 09AGBI PRELIMINARY; PRT; 165 AA.
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OSCP-MODIFIED.
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB19;
 RA Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
 RA Lawson C.L., Swaminathan S.;
 RT "Crystal structure of outer surface protein C (OspC) from the Lyme
 RT disease spirochete, Borrelia burgdorferi.";
 RL EMBO J. 0:0-0(2001)
 DR EMBL; AF337548; AAK21289.1; -
 FT VARIANT 61 61 M -> I.
 FT SEQUENCE 165 AA; 17833 MW; 56E01536D22F61BF CRC64;
 SQ

Query Match 42.7%; Score 786; DB 2; Length 165;
 Best Local Similarity 98.2%; Pred. No. 3.7e-29;
 Matches 161; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

212 VGPMLTEISKRTESNAVAVLAVKEVEALLSSIDEIAKIGKKIRNDVSLDNEADHNSL 271
 1 MKGPMLTEISKRTESNAVAVLAVKEVEALLSSIDEIAKIGKKIRNDVSLDNEADHNSL 60

QY 272 ISGAYIISMLITKKISATDSGELKAELKAKKCSSEFTAKLGEHTDLGKEGYTDDNAK 331
 DB 61 MSGAYIISMLITKKISATDSGELKAELKAKKCSSEFTAKLGEHTDLGKEGYTDDNAK 120

QY 332 KAILKTNDKTGADELKLFESVKNLSKAKEMLTNSVKELTS 375
 DB 121 KAILKTNDKTGADELKLFESVKNLSKAKEMLTNSVKELTS 164

RESULT 15
 ID 09R7B0 PRELIMINARY; PRT; 159 AA.
 AC 09R7B0:
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;

RX MEDLINE-97478003; PubMed-9336916;
 RA Ras N.M., Postic D., Foretz M., Baranton G.;
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
 RT U.S.A.'?";
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;
 RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
 RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U91799; AA881896.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 159
 FT NON_TER 1 159
 FT SEQUENCE 159 AA; 16861 MW; 5285D422C43DC213 CRC64;
 SQ

Query Match 42.0%; Score 773; DB 2; Length 159;
 Best Local Similarity 99.4%; Pred. No. 1.4e-28;
 Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 TSANADESVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAKIGKKIHONGNL 71
 DB 1 TSANADESVKGPMLTEISKRTIDSNVAVLAVKEVEALLSSIDEIAKIGKKIHONGNL 60

QY 72 DTEYNHNSGLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFTTNKLEKHTDLGK 131
 DB 61 DTEYNHNSGLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFTTNKLEKHTDLGK 120

QY 132 EGYTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKA 170
 DB 121 EGYTDADAKAELIKTNGTKTGAEELGKLFESVEVLSKA 159

Search completed: March 18, 2002, 10:10:53
 Job time: 976 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:42 ; Search time 39.62 Seconds
(without alignments)
347.030 Million cell updates/sec

Title: US-09-596-746a-24
Perfect score: 1842

Sequence: 1 MACNNSGDCNTSANSADSE.....KNLSRAKEMLTNSVKELTS 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

Number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	50.4	210	1	OSCL_BORBU
2	614	33.3	212	1	OSCL_BORBU
3	404.5	22.0	214	1	VM24_BORHE
4	400	21.7	215	1	VM03_BORHE
5	176.5	9.6	1509	1	MYSN_ACAKA
6	170	9.2	1251	1	RBP2_PLAVB
7	169	9.2	1957	1	YD86_SCHPO
8	167	9.1	483	1	M6_STRPY
9	167	9.1	564	1	M12_STRPY
10	163.5	8.9	1790	1	USOL_YEAST
11	162.5	8.8	3210	1	CENF_HUMAN
12	159	8.6	2116	1	MYSS_DICDI
13	157.5	8.6	1786	1	LMBI_MOUSE
14	157	8.5	998	1	SCAA_RICAK
15	157	8.5	1938	1	MYN_AEOIR
16	155	8.4	705	1	YNPQ_CAEEL
17	154	8.4	1966	1	MYSE_CAEEL
18	153	8.3	1427	1	REST_HUMAN
19	152	8.3	492	1	M5_STRPY
20	150.5	8.2	775	1	YHGE_BACSU
21	148.5	8.1	1713	1	LMAS_HUMAN
22	148	8.0	1433	1	REST_CHICK
23	147	8.0	1164	1	BAG_STRAG
24	146.5	8.0	875	1	ZIP1_YEAST
25	146	7.9	1786	1	LMBI_HUMAN
26	146	7.9	2349	1	TPR_HUMAN
27	145.5	7.9	1500	1	MYSS_STRGN
28	145.5	7.9	1961	1	MYH5_RAT
29	145	7.9	444	1	PUR2_MERJA
30	145	7.9	962	1	VDP_HUMAN
31	145	7.9	2869	1	RBP1_PLAVB
32	145	7.9	3110	1	LMAS_HUMAN
33	144	7.8	1189	1	SCIL_CHICK

ALIGNMENTS

RESULT ID	1	OSCL_BORBU	STANDARD:	PRT:	210 AA.
AC	007337				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB19.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Plasmid 1p54.				
OX	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
NCBI_Taxid:	139;				
PN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-93239332; PubMed-8098841;				
RA	Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,				
RA	Will G., Wilske B.;				
RT	"Genetic heterogeneity of the genes coding for the outer surface				
RT	protein C (ospc) and the flagellin of Borrelia burgdorferi.";				
RL	Med. Microbiol. Immunol. 182:37-50(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-94041630; PubMed-8225587;				
RA	Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;				
RT	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-96025162; PubMed-7494039;				
RA	Fukunaga M., Hamase A.;				
RT	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan.";				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-98065943; PubMed-9403685;				
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kertlavage A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,				

099323 drosophila
P34237 saccharomyc
060675 mus musculu
P02362 oryctolagus
090339 cypripinus ca
P41541 bos taurus
094181 rickettsia
P51834 bacillus su
091783 xenopus lae
P12753 saccharomyc
P28025 xenopus lae
Q9516 schizosacch

RA Utterback T., Matthey L., McDonald L., Artlach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi";
 RL Nature 390:580-586(1997).
 CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X69596; CAA49306.1; -
 CC EMBL: U01894; AAA16058.1; -
 CC EMBL: D49487; BAA08457.1; -
 CC EMBL: AE00792; AAC66329.1; -
 CC TIGR: BBH19; -
 CC InterPro: IPR001800; Lipoprotein_6.
 CC Pfam: PF01441; Lipoprotein_6; 1.
 CC Prodom: PD001149; Lipoprotein_6; 1.
 CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
 CC Complete proteome.
 CC KW SIGNAL 1 18 BY SIMILARITY.
 CC FT CHAIN 19 210 OUTER SURFACE PROTEIN C.
 CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
 CC SQ SEQUENCE 210 AA; 22340 MW; 7A4FC97BF9177BF CRC64;
 CC -----
 CC Query Match 50.4%; Score 928; DB 1; Length 210;
 CC Best Local Similarity 98.4%; Pred. No. 1.3e-35;
 CC Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC 1 MACNNSGKDGNTSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 60
 CC ::
 CC 17 ISCNNSGKDGNTSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 76
 CC 61 IGKRIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNEGLKRTDAKKCSEFTN 120
 CC ::
 CC 77 IGKRIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNEGLKRTDAKKCSEFTN 136
 CC 121 KLEKHTDLDGREGVTDADAKKALIKTNGTKGAELIGKLFESVEVLSKAKEMLANSYK 180
 CC ::
 CC 137 KLEKHTDLDGREGVTDADAKKALIKTNGTKGAELIGKLFESVEVLSKAKEMLANSYK 196
 CC -----
 CC 181 ELTSPVAESP 191
 CC ::
 CC 197 ELTSPVAESP 207
 CC -----
 CC RESULT 2
 CC OSC2_BORBU STANDARD; PRT; 212 AA.
 CC AC 008137;
 CC DT 15-DEC-1998 (rel. 37, Created)
 CC DT 15-DEC-1998 (rel. 37, Last sequence update)
 CC DT 15-DEC-1998 (rel. 37, Last annotation update)
 CC DE OUTER SURFACE PROTEIN C PRECURSOR (PC).
 CC OSpc.
 CC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 CC NCBI_TaxID=139;
 CC [1]
 CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC STRAIN=PKO;
 CC MEDLINE=92219955; PubMed=1560779;
 CC Fuchs R., Jauris S., Lottspeich F., Praeger-Mursic V., Wilske B.,
 RA

RA Soutschek E.;
 RT "Molecular analysis and expression of a Borrelia burgdorferi gene
 RT encoding a 22 kDa protein (pc) in Escherichia coli";
 RL Mol. Microbiol. 6:503-509(1992).
 CC [2]
 CC SEQUENCE OF 1-205 FROM N.A.
 CC STRAIN=DK26;
 CC MEDLINE=94075528; PubMed=8253951;
 CC RA Theisen M., Frederiksen B., Lebesch A.M., Vuust J., Hansen K.;
 RT "Polymorphism in ospC gene of Borrelia burgdorferi and
 RT immunoreactivity of ospC protein: implications for taxonomy and for
 RT use of ospC protein as a diagnostic antigen";
 RL J. Clin. Microbiol. 31:2570-2576(1993).
 CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X62162; CAA44093.1; -
 CC EMBL: X73624; CAA52003.1; -
 CC InterPro: IPR001800; Lipoprotein_6.
 CC Pfam: PF01441; Lipoprotein_6; 1.
 CC Prodom: PD001149; Lipoprotein_6; 1.
 CC Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
 CC KW SIGNAL 1 18 BY SIMILARITY.
 CC FT CHAIN 19 212 OUTER SURFACE PROTEIN C.
 CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
 CC SQ SEQUENCE 212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;
 CC -----
 CC Query Match 33.3%; Score 614; DB 1; Length 212;
 CC Best Local Similarity 69.9%; Pred. No. 1.4e-21;
 CC Matches 135; Conservative 14; Mismatches 42; Indels 2; Gaps 2;
 CC -----
 CC 1 MACNNSGKDGNTSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 59
 CC ::
 CC 17 ISCNNSGKDGNTSANSADSVKGNLTETSKRTTDSNAVLAVKEVALLSIDEIAAK 76
 CC 60 AIGKRIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNEGLKRTDAKKCSEFT 118
 CC ::
 CC 77 AIGKRIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNEGLKRTDAKKCSEFT 136
 CC 119 TNKLEKHTDLDGREGVTDADAKKALIKTNGTKGAELIGKLFESVEVLSKAKEMLANS 178
 CC ::
 CC 137 TNKLEKHTDLDGREGVTDADAKKALIKTNGTKGAELIGKLFESVEVLSKAKEMLANS 196
 CC -----
 CC 179 VKELTSPVAESP 191
 CC ::
 CC 197 VKELTSPVAESP 209
 CC -----
 CC RESULT 3
 CC VM24_BORHE STANDARD; PRT; 214 AA.
 CC AC P32778;
 CC DT 01-OCT-1993 (rel. 27, Created)
 CC DT 01-OCT-1993 (rel. 27, Last sequence update)
 CC DT 01-FEB-1994 (rel. 28, Last annotation update)
 CC DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
 CC VMP24.
 CC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 CC NCBI_TaxID=140;
 CC [1]
 CC SEQUENCE FROM N.A.
 RA

```

RC STRAIN-SSP, HSI SEROTYPE 24;
RX MEDLINE-93133110; PubMed-1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids
are highly polymorphic.";
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).
-----
CC EMBL: L04786; AAA22964.1; -
CC InterPro: IPR001800; Lipoprotein_6.
CC Pfam: PF01441; Lipoprotein_6; 1.
CC ProDom: PD001149; Lipoprotein_6; 1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Outer membrane; Lipoprotein; Signal; Plasmid.
CC CHAIN 1 18 PROBABLE.
CC STAGNAL 1 18 VARIABLE MAJOR OUTER MEMBRANE
CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
CC SO SEQUENCE 214 AA; 22541 MW; F1583F510246FC7 CRC64;

Query Match 22.0%; Score 404.5; DB 1; Length 214;
Best Local Similarity 45.4%; Pred. No. 3.1e-12;
Matches 93; Conservative 36; Mismatches 59; Indels 17; Gaps 5;

QY 1 MACNSNGKNGTNSANSADSVKGP----NTEISKTKTDSNAVLAVKEVALLSSIDEI 56
DB 17 MSCNNGSGPE-----LKSDEVAKSDGTVDLAKVSKIKIKESAPASVKEVETLVKSYDEL 71
QY 57 AAKAIGKKI-HONNGDTEVNHNGSLAGAVAI STLKOKL-----DGLNKGLEKIDA 110
DB 72 -AAAIKIKIKINDGDLDTENGONGSLAGVSVSAVKIKGALETTSGISINE-LKTITE 129
QY 111 AKKCSFTFNKLEKHTDLGKEGVTDAKAEALIKTNGTKGAELGKLFESYEVLKA 170
DB 130 VKSKAEAFLEKLTGDKGTTELCKKIDSDDTKKAIKKIDNSDKGASELEALNTAVDALLKA 189
QY 171 AKEMLANSVKELTSPVAESPAMVN 195
DB 190 AEGVEEAIAIKELTAPVAKERPSQNN 214

RESULT 4
VMO3_BORHE STANDARD; PRT; 215 AA.
AC 002448;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
OS VMP3.
GN Borrelia hermsli.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBL_Taxid-140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SSP, HSI SEROTYPE 3;
RX MEDLINE-93133110; PubMed-1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

```

```

RT are highly polymorphic.";
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP24.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).
-----
CC EMBL: L04789; AAA22967.1; -
CC InterPro: IPR001800; Lipoprotein_6.
CC Pfam: PF01441; Lipoprotein_6; 1.
CC ProDom: PD001149; Lipoprotein_6; 1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Outer membrane; Lipoprotein; Signal; Plasmid.
CC CHAIN 1 18 PROBABLE.
CC STAGNAL 1 18 VARIABLE MAJOR OUTER MEMBRANE
CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
CC SO SEQUENCE 215 AA; 23139 MW; 684C74D35F87C771 CRC64;

Query Match 21.7%; Score 400; DB 1; Length 215;
Best Local Similarity 45.9%; Pred. No. 5e-12;
Matches 94; Conservative 34; Mismatches 61; Indels 16; Gaps 5;

QY 1 MACNSNGKNGTNSANSADSVKGP----NTEISKTKTDSNAVLAVKEVALLSSIDEI 56
DB 17 MSCNNGSGPE-----LKSDEVAKSDGTVDLAKVSKIKIKESAPASVKEVETLVKSYDEL 71
QY 57 AAKAIGKKI-HONNGDTEVNHNGSLAGAVAI STLKOKL-----NEGLEKIDA 110
DB 72 -AAAIKIKIKINDSNEFEDNDHNSLAGVGVILTYAKATLSLEQITIGISDELKTEYGM 130
QY 111 AKKCSFTFNKLEKHTDLGKEGVTDAKAEALIKTNGTKGAELGKLFESYEVLKA 170
DB 131 VKKSEAFVGVQVSKKHNDLAKEGVTDAHAASALLVDTGTXKGAELIKLMTAIDELKA 190
QY 171 AKEMLANSVKELTSPVAESPAMVN 195
DB 191 ANDAVETVIKELTASVAKERPSQNN 215

RESULT 5
MYSN_ACACA STANDARD; PRT; 1509 AA.
AC P05659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS Acanthamoeba castellanii (Amoeba).
CC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBL_Taxid-5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87308395; PubMed-3040773;
RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail.";
RL J. Cell Biol. 105:913-923(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
INFO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

```

CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (MLC-2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
 CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: Y00624; CAA68663.1; .
 DR PIR: A27224; A27224.
 DR HSSP: P08799; IAMD.
 DR InterPro: IPR000048; IQ.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 KW Myosin; coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
 KW Methylation; Alkylation; Phosphorylation; Multigene family.
 FT DOMAIN 1 789
 FT DOMAIN 790 819
 FT DOMAIN 848 1509
 FT DOMAIN 848 1226
 FT DOMAIN 1227 1252
 FT DOMAIN 1253 1509
 FT DOMAIN 1253 1482
 FT DOMAIN 1483 1509
 FT NP_BIND 182 189
 FT DOMAIN 660 682
 FT DOMAIN 766 780
 FT MOD_RES 133 133
 FT MOD_RES 700 700
 FT MOD_RES 1489 1489
 FT MOD_RES 1494 1494
 FT MOD_RES 1499 1499
 FT MOD_RES 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;
 SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;
 Query Match 9.6%; Score 176.5; DB 1; Length 1509;
 Best Local Similarity 24.8%; Pred. No. 0.41;
 Matches 103; Conservative 68; Mismatches 137; Indels 107; Gaps 21;
 Oy 28 ELSKRTDSNAVLAVKEVEA-----LLSSIDE-----IAAKIGKTHQNGDTEYNH 77
 Db 934 ELQETTSASNDILEQKRLLEAKGELKASLEEEENRKALQDEAKTVESSENELODYED 993
 Oy 78 NCSLLAGYAISTLIKOKLDLKGKLEKEKIDA--AKKCEFTFNKLEKHTDLGKEGV 134
 Db 994 EA-----AAHDSLKKEEDLSRE-LRETKDALDAENISETLRSKL--KNTERGADVV 1043
 Oy 135 TDA--DAKEALIKTNGTGTGAELGKLFESVEVLSKAEMLANSYE----- 181
 Db 1044 RNEEDDVATATLOLEKTKKSLLEELAQTRQLEE-EKSGKAASSKAKKOLGOOLEDAARSE 1102
 Oy 182 ---LTSPYVAESPAMVNSGKDN-----TSANSADESVKG--DNLEISKTKIT 225
 Db 1103 VDSIKSKLSIAEKSL--KTAKDNQRDLEQDLEDERTVRAVNDOKKALEKTELEDOVT 1160

Oy 226 ---ESNAVVLAVKEVETLTSIDEL-----AKAIGKIKNDVSLDNEADHNSGL 271
 Db 1161 ALGQOKKA---AAQAKTLTQYDETKRRLLEAEASAAERKERKN--ALDEVAQ----- 1210
 Oy 272 ISGAYLISNLITTKISAIDSG-----ELKAEIERAKK---CSEFTKLKGE 316
 Db 1211 -----LTADLDAERDSGAQORRKLNTRISELSQSLFNAPKGGASSEVVRLEGE 1260
 Oy 317 HTDLGKEGVDDAKKAILKTNNDKTKGADLEKLFESYKNSKAKEMLTNSVK 771
 Db 1261 LELEBELLLTAQRAKAAK--NIDKAN--LELEELQEAADARNDKLVKDNKR 1312
 RESULT 6
 RBP2_PLAVB STANDARD; PRT; 1251 AA.
 ID RBP2_PLAVB
 AC 000799;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
 GN RBP2.
 OS Plasmodium vivax (strain Belen).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=31273;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92315338; PubMed=1617731;
 RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
 RT "A reticulocyte-binding protein complex of Plasmodium vivax
 RT merozoites".
 RL Cell 69:1213-1226(1992).
 CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
 CC HUMAN RETICULOCYTE CELLS.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M88098; AAA29744.1; .
 KW Malaria; Receptor; Membrane.
 FT NON_TER 1
 FT NON_TER 1251
 FT NON_TER 1251
 SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;
 Query Match 9.2%; Score 170; DB 1; Length 1251;
 Best Local Similarity 22.9%; Pred. No. 0.65;
 Matches 93; Conservative 50; Mismatches 139; Indels 124; Gaps 17;
 Oy 22 KGNLTETSKRTDSNAVLAVKEVEALLSSIDEIYAKA--IGKTHQNGDTEYNHG 79
 Db 774 KNTNEEDVHKNQDAVKVLEI-----LAHDEIDTKKDSKSLIEMNQI----- 819
 Oy 80 SLLAGYAISTLIKOKLDLKGKLEKEKIDA--AKKCEFTFNKLEKHTDLGKEGV 136
 Db 820 -----YKRVLLINQ-----YKKNISIKSKEEAVSVKIGVNSKSHSELKITSQSD 864
 Oy 137 ADAKEAIL-----KTNKTKTGAELGKLFESVEVLSKAEMLANSY 179
 Db 865 KSYNDIALLEKQTELONLNSPTQEKNTNSKLEKTKTDES-----LKNAL 913
 Oy 180 KELTSYVAESPAMVNSGKDNSTANSADSVKGNPLTETSKRTITSENAVLAVKEVET 239
 Db 914 KTLGEVNA-----LKASDNEHGVQSKSPV--NPALSEIEKEBT-----DIDS 956
 Oy 240 LITSIDELAK-----AIGKIKNDVSLDNEADHNSGLISGAYL----- 277

```

Db 957 LNTALDELLKGRGTCVSRKILKDTVTKEISDDELINTIEKNV----AYLAITKKNV 1012
Qy 278 ---ISLITL--KTSIAIKDSGELKAEIERAKKCSSEPTAKLGEHTDLGEGVTDNAK 332
Db 1013 EDTVOOVLTNLEHFNTRKQVSNHEPTNFKSNKSEELTKAVTDSKITISK-----LKG 1065
Qy 333 ALIKTN-----NDKTGKADELEKLFESVKNLSAAKEM--LTNSKV 371
Db 1066 VIEENVNEMNTIESSAKEITELNLEKKNKTSLTNEITOTSNEVK 1111

RESULT 7
YD86_SCHPO STANDARD: PRT; 1957 AA.
ID YD86_SCHPO
AC Q10411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN SPAC1F3.06C.
Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 270690; CAA94624.1; -.
DR Hypothetical protein.
KW SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;
SQ

```

```

Db 690 NIOTLEKEDLKKSEELAFKFSLEAKNLEFVIDNLKGKHEHTLEAQNDHSLSDAKKNTNAI 749
Qy 335 LKTNNDKTKGADELEKLFESVKNL---SKAKEMLTNSVELTS 375
Db 750 L--SSLETKSESDVKRLTANVELTLODSKAMKQSFSTLVNSYGS 791

RESULT 8
M6_STRPY STANDARD: PRT; 483 AA.
ID M6_STRPY
AC P08089;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE M PROTEIN, SEROTYPE 6 PRECURSOR.
GN EMM6.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacilllus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86111835; PubMed=3511046;
RA Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT "Complete nucleotide sequence of type 6 M protein of the group A
RT Streptococcus. Repetitive structure and membrane anchor.";
RL J. Biol. Chem. 261:1677-1686(1986).
RN [2]
RP SEQUENCE OF 43-122 FROM N.A.
RX MEDLINE=8516224; PubMed=3885219;
RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT "Relationship of M protein genes in group A streptococci.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC PHAGOCYTOSIS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -1- SIMILARITY: TO OTHER M PROTEINS.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M1138; AAA26920.1; -.
DR PIR; A26297.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PR00015; GP05ANCHOR.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW Transmembrane; Coiled coil; Signal.
FT SIGNAL 1 42
FT CHAIN 43 483
FT DOMAIN 43 457
FT TRANSMEM 458 477
FT DOMAIN 478 483
FT DOMAIN 478 483
FT DOMAIN 69 138
FT DOMAIN 157 269
FT DOMAIN 279 347
FT DOMAIN 348 411
FT DOMAIN 412 448
FT DOMAIN 449 454

```

FT PROTEINS.
SQ SEQUENCE 483 AA: 53472 MW: 6687FF28DB53AA48 CRC64;

Query Match 9.1%; Score 167; DB 1; Length 483;
Best Local Similarity 23.3%; Pred. No. 0.31;
Matches 91; Conservative 57; Mismatches 157; Indels 86; Gaps 14;

QY 5 NSGADGNTSANSADSVKGNLFEISKRTDSNAVLAVVEVALLSIDELAKAIGK 64
DB 73 NNDKLTENNLDON---KNLTENNLDONKNLTENN---NL 112
QY 65 IHONNGDTEYNHNGSLAGAYASTLIKQKLDGKNEGKEKIDAKKSEFTFNK 124
DB 113 TDQKNLTTE---NKEIKAEENRLT-----EKKGILTKLSEF---EBAKERE 157
QY 125 KHHDLGK-EGVTDADAEALIKNGTKTCAEELGKLFESVE---VLSKAKE- 177
DB 158 NKEAIGTKLTDETVKDKLAKQESK---ETIGTLKLTDETVKDKLAKQESKETIG 213
QY 178 SVKELTSPVVAESPAMVNNNGKDGNTSANSADSVKGNLFEISKRTDSNAVLAVKEV 237
DB 214 TLKKTLDTEYKDKIAKEQESKETIGTLKLTDETVK---DKTARE-QSKQDIGALKOE 268
QY 238 ETLTSLDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISMLTKRTISAIDSGELKA 297
DB 269 LAKKDEGNKVSSEASRKGRLRDLASREA-----KKQVEKDLANLTA 309
QY 298 EIEKAKKSEFTFAKLGERTDGLKGEVDDNNAKKAILKTNDKGTGADELEKLFESVKN 357
DB 310 ELDKVKEKQISDASKQGLRDLASREAKQVEKA-LEBANSKLALEKLEKLEESK 368
QY 358 L-----SKAKEMLTNSVKEL 373
DB 369 LFEKERAELOAKLEAKLEAKLEQAKQAEEL 399

RESULT 9
M12_STRPY STANDARD; PRT; 564 AA.

AC P19401;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).
GN EMM12.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;

RP SEQUENCE FROM N.A.
RC STRAIN-CS24 / Serotype M12;
RA MEDLINE=88058777; PubMed=2445730;
RA Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;
RT "Streptococcus pyogenes type 12 M protein gene regulation by upstream sequences";
RL J. Bacteriol. 169:5633-5640(1987).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO PHAGOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/>).

CC or send an email to license@isb.ch.
CC -----
DR EMBL: M1269; AAA88573.1; -
DR PIR: A60115; A60115.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF00746; Gram_pos_anchor. 1.
DR Pfam: PF02370; M: 9.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen; Transmembrane; Coiled coil; Signal.
FT SIGNAL 1 41
FT CHAIN 42 564 M PROTEIN, SEROTYPE 12.
FT DOMAIN 42 550 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 551 >564 MEMBRANE ANCHOR.
FT DOMAIN 44 505 COILED COIL (POTENTIAL).
FT DOMAIN 505 541 GLY/PRO-RICH (CELL WALL-SPANNING).
FT DOMAIN 542 547 CONSERVED IN GRAM-POSITIVE COCCI SURFACE PROTEINS.
FT NON_TER 564 564
FT SQ SEQUENCE 564 AA: 62904 MW: 5F1549DACAA77B46 CRC64;

Query Match 9.1%; Score 167; DB 1; Length 564;
Best Local Similarity 22.6%; Pred. No. 0.36; Mismatches 161; Indels 116; Gaps 17;
Matches 97; Conservative 55; Mismatches 157;

QY 10 GNTSANSADSVKGNLFEISKRTDSNAVLAV-KEYEALLSIDELAKAIGKIKHON 68
DB 115 GKLIGIDNLDLAKK---ITELEKSYVEKNDVLSQIKKELEAEKQIDQ-----FGREYHAA 165
QY 69 NGLDTEYNHNGSLAGAYASTLIKQKLDGKNEGKEKIDAKKSEFTFNKLEKHTD 128
DB 166 DIL-----RKQELIAKEKNVIS-----KLNG-ELQPLKQKVD-----ETDRNLQOE----- 205
QY 129 LKGEVTDADAEALIKNGTKTCAEELGKLFESVEVLSRAA--KEMIANSVKELTSPV 186
DB 206 --KQKVLSEQOLAVTKENAKKDFELALGHQLDKEYNAKIAELESKLADAKKDFELAA 263
QY 167 V-----AESPMVNNNGKGC-----NTSANSADSVKGC-----NL 218
DB 264 LGHOHAHEYQAKLEKQKQIKOLEEKOILDASKRGATRDLEAVRAKKAETAEELNNLK 323
QY 219 EISKRTESNAVLAVK-----EVELLTSTIDE---LAKAIGKIKNDV 259
DB 324 AELAKVTEQKQIILDASRGATRDLEAVRKKSKQVEALAKLEQONNISEASRKGRLDL 383
QY 260 SLDNEADHNGSLISGAYLISMLTKRTISAIDSGELAEIEKAKKSEFTFAKLGERTD 319
DB 384 DTSREA-----KKQVEKDLANLTAELDKVKEEKQISDASRQGLRD 424
QY 320 LKGEVTDNNAKKAILKTNDKGTGADELEKLFESVKNL-----SKAKE 364
DB 425 LDASREAKQVEKA-LEBANSKLALEKLNKDNDEESKRLTEKEKAELOAKLEAKALKE 483
QY 365 MLTNSVKEL 373
DB 484 QLAQAEEL 492

RESULT 10
USOL_YEAST STANDARD; PRT; 1790 AA.

AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.
GN USOL OR INT1 OR YDI058W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-X2180-1A;
 RX MEDLINE-91183402; PubMed-2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 RA Yamasaki M.;
 RT "A cytoskeleton-related gene, usol, is required for intracellular
 RT protein transport in *Saccharomyces cerevisiae*.";
 RL J. Cell Biol. 113:245-260(1991).
 RN [2]
 RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hosetter M.K., Herman D.J., Bendel C.M., McCellan M., Tao N.,
 RA Kendrick K.E.;
 RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Bai Y., Symington L.S.;
 RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
 CC ER AND THE GOLGI COMPLEX.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
 CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X54378; CA38253.1; -
 DR EMBL: L03188; AAB00143.1; -
 DR EMBL: U53668; AAB6659.1; -
 DR PIR: A38455; A38455.
 DR HSSP: P80220; 1DIP.
 DR SGD: S0002216; USO1.
 DR Interpro: IPR002017; Spectrin.
 KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
 FT DOMAIN 1 724 GLOBULAR HEAD.
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
 FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
 FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
 FT CONFLICT 847 847 G->E (IN REF. 2).
 FT CONFLICT 924 924 E->K (IN REF. 2).
 FT CONFLICT 1253 1253 V->I (IN REF. 2).
 FT CONFLICT 1319 1319 N->V (IN REF. 2).
 FT CONFLICT 1461 1461 G->S (IN REF. 2).
 FT CONFLICT 1581 1581 G->S (IN REF. 2).
 FT CONFLICT 1600 1600 I->V (IN REF. 2).
 FT CONFLICT 1661 1661 R->S (IN REF. 2).
 FT CONFLICT 1772 1772 D->DEEDDEE (IN REF. 2).
 SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;
 Query Match 8.9%; Score 163.5; DB 1; Length 1790;
 Best Local Similarity 21.8%; Pred. No. 1.9;
 Matches 102; Conservative 73; Mismatches 184; Indels 109; Gaps 17;
 OY 3 CUNNSGKDG-----MTSANSADSVGPNLTETSKITTSNANVLAVKFEVALSS 52
 DB 929 CUNLSKEKHISLVEYKSRFOSHNLY--AKLTETKLSLANNYKDMA--ENESLTKA 984
 OY 53 IDEIAKAIGKRIHONNGLD--TEYNHNGSLLAGAY-----AISTL-----IKOK 95
 DB 985 VEESKNESSIQLSNLQNKIDMSQOEKENFOIERGSIETKNIETQLKTTISDLEQTKREELISK 1044

OY 96 LDGLKNGC-----LKEIDAAKKSEFTFNKLE-----KHTDLGKEGVTD 136
 DB 1045 SDSSKDEYESQISLKEKLELTATTANDENAVNKISELTTRTEELAEALAKNLKNELET 1104
 OY 137 ADAREKALK-----TNGTKTK-----GAEELGKLFESVEVLSKAA 171
 DB 1105 LEISEKALKVEKENEHLKREKIQLEKEATREYQOQLNSLANIESLEKHEHDLAAQKKY 1164
 OY 172 KEMLAANSVKELTSPV-----VAESPANVNSGKDGTANSADSVGPNLTETSKITE 226
 DB 1165 EEOIANKEROYNEEISOLNDEITSTOQENESIKKKNDLGEVYKAMKSTSEOSNLKSE 1224
 OY 227 SNAVVLAVKREV-----ETLNSIDELAKAIGKRIKNDVSLNDEADHNSLSGAVLISNLI 282
 DB 1225 IDALNTQIKELKKNENENESLSIESISVESEYTKIKELDDECFKEKVESE-----L 1277
 OY 283 TKRISAIKIDSG-----ELKAEIERAKKCEFTAKL-----GEHTDLGKEGVTDNAKRAIL 335
 DB 1278 EDKLKASEDKNSKYLDELQKSEKIKELDAKTTELKIQLEKITLNSAKKESSELSRLK 1337
 OY 336 KTNNDKTKGADE-LEKL-----FESYKNSLKAKEMLNSVKE 372
 DB 1338 KTSSEERKNAEDELKELKNEIQIKNOAFERKERKILNKGSSITTOEYSE 1385
 RESULT 11
 ID CENF_HUMAN STANDARD; PRT; 3210 AA.
 AC P49454; Q13246; Q13171;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 20-FEB-1996 (Rel. 33, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CENP-F KINETOCHORE PROTEIN (CENTROMERE PROTEIN F) (MITOSIN) (AH
 DE ANTIGEN).
 GN CENPF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Breast carcinoma;
 RX MEDLINE-95348175; PubMed-7542657;
 RA Liao H., Winkfein R.J., Mack G., Rattner J.B., Yen T.J.;
 RT "CENP-F is a protein of the nuclear matrix that assembles onto
 RT kinetochores at late G2 and is rapidly degraded after mitosis.";
 RL J. Cell Biol. 130:507-518(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95379848; PubMed-7651420;
 RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B.,
 RA Jones D., Yang-Feng T.L., Lee W.-H.;
 RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein
 RT that is specifically involved in mitotic-phase progression.";
 RL Mol. Cell. Biol. 15:5017-5029(1995).
 RN [3]
 RP SEQUENCE OF 2194-3210 FROM N.A.
 RX MEDLINE-9536446; PubMed-7612011;
 RA Li Q., Ke Y., Kapp J.A., Fertig N., Medsker T.A. Jr., Joshi H.C.;
 RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal
 RT domain sufficient for nuclear localization.";
 RL Biochem. Biophys. Res. Commun. 212:220-228(1995).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE-95370296; PubMed-7642639;
 RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;
 RT "The C terminus of mitosis is essential for its nuclear localization,
 RT centromere/kinetochore targeting, and dimerization.";
 RL J. Biol. Chem. 270:19545-19550(1995).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE-98437347; PubMed-9763420;
 RA Chan G.K.T., Schaar B.T., Yen T.J.;

RT "Characterization of the kinetochore binding domain of CENP-E reveals
 RT interactions with the kinetochore proteins CENP-F and hBUBR1.";
 RT J. Cell Biol. 143:49-63(1998).
 CC -1- FUNCTION: PROBABLY REQUIRED FOR KINETOCORE FUNCTION, INVOLVED IN
 CC CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH
 CC KINETOCORE PROTEIN (KCP), CENP-E AND BUBR1.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),
 CC REORGANIZATION TO THE KINETOCORE/CENTROMERE (CORONAL SURFACE OF
 CC THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.
 CC -1- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.
 CC -1- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: U19769; AAA82889.1; -;
 DR EMBL: U30872; AAA82935.1; -;
 DR EMBL: U25725; AAA86889.1; -;
 DR HSSP: P02649; 1LE4.
 DR MIM: 600236; -;
 DR
 KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil;
 KM Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.
 FT DOMAIN 14
 FT DOMAIN 197
 FT DOMAIN 273 769
 FT DOMAIN 823 1328
 FT DOMAIN 1642 1746
 FT DOMAIN 1862 2987
 FT DOMAIN 2207 2568
 FT REPEAT 2207 2386
 FT REPEAT 2389 2568
 FT REPEAT 3015 3032
 FT SITE 16 16
 FT CONFLICT 250 250
 FT CONFLICT 272 272
 FT CONFLICT 611 611
 FT CONFLICT 1494 1589
 FT CONFLICT 1611 1611
 FT CONFLICT 1811 1811
 FT CONFLICT 2242 2243
 FT CONFLICT 2335 2335
 FT CONFLICT 2492 2492
 FT CONFLICT 2545 2361
 SQ SEQUENCE 3210 AA; 367589 MW; 11D83324960E4334 CRC64;
 Query Match 8 8%; Score 162.5; DB 1; Length 3210;
 Best Local Similarity 22.0%; Pred. No. 4;
 Matches 84; Conservative 56; Mismatches 147; Indels 95; Gaps 12;
 Oy 44 KEVALLSSIDIDIAKAG--KRIHONNGDTEVHNHNSLLAGVATSTLKOKDLSK 100
 Db 2208 EEVHQLRGIEKLRVRIEADKKQIHLKLERRENDL-----KQKVENLE 2256
 Oy 101 N-----EGLKEKIDAAKCEFTFNKLEKHTDQKGVTDADAKAHL 144
 Db 2257 RELQSENOELVILDAENSKAEVETLTKQJEEMARSLKVFELD---VLRSEKELUT 2312
 Oy 145 KTNNGTKGAEELKLFESVYL-----SKAKEMANSVKELTSPVAES 190
 Db 2313 KOIQEKQGLSELBKLISFSLSEKQADIQIKESKTAVENTMQLKELNEAV---- 2368
 Oy 191 PAMVNSGKDNSTANSADSVKGNPLEISKTTESNAVLAKEVEYTLTSTDELAKA 250
 Db 2369 -----AALCGDQIKMATQSDLPPEIEHQLRNSIEKIRARLEADKKQOLC 2415
 Oy 251 IGGIKKNDVSLDNEADHNGSLISGAVYLSNL-----ITKKISAIKDSGELKAEIEKAK 303

Db 2416 VLOOLK-----ESEHADLLKGR--VENLELEIARTNOEHALENSKGEVETLK 2466
 Oy 304 KCEEFKAKLKGHTL-----GKEGVTDNAK-----ALKTNDKYGAELEKL 351
 Db 2467 AKLEGKQSLRGLELDVYIRSEKEDLTNLEKQERISELEINSSFENIQERQEKV 2526
 Oy 352 FESVKNLSKAKEMLTNSVKEL 373
 Db 2527 -OMKEKSSYAMEMLOQLKEL 2546
 RESULT 12
 MS2.DICDI
 ID MS2.DICDI STANDARD; PRT; 2116 AA.
 AC P08799;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
 GN MICA.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87092266; PubMed=3540939;
 RA Warrick H.M., de Lozanne A., Leitwand L.A., Spudich J.A.;
 RT "Conserved protein domains in a myosin heavy chain gene from
 RL Dictyostelium discoideum.";
 RN Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
 RP [2]
 RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
 RC STRAIN=AX2;
 RX MEDLINE=90353583; PubMed=2387408;
 RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
 RT Gerisch G.;
 RL "Replacement of threonine residues by serine and alanine in a
 RT phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
 RN FEBS Lett. 269:239-243(1990).
 RP [3]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=88112226; PubMed=2828113;
 RA Magle G., Noegel A., Scheel J., Gerisch G.;
 RT "Phosphorylation of threonine residues on cloned fragments of the
 RL Dictyostelium myosin heavy chain.";
 RN FEBS Lett. 227:71-75(1988).
 RP [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345066; PubMed=7619795;
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
 RT Rayment I.;
 RL "X-ray structures of the myosin motor domain of Dictyostelium
 RT biochemically complexed with MgADP.Berx and MgADP.ALf4-";
 RN Biochemistry 34:8960-8972(1995).
 RP [5]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345067; PubMed=7619796;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
 RL truncated head of Dictyostelium discoideum myosin to 2.7-A
 RT resolution.";
 RN Biochemistry 34:8973-8981(1995).
 RP [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=96206189; PubMed=8611550;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
 RL Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
 RN Biochemistry 35:5404-5417(1996).
 RP [7]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
 RX MEDLINE=97452580; PubMed=9305951;

RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
 RT "X-ray structures of the MgADP, MgATPgammAS, and MgAMPPNP complexes
 of the Dictyostelium discoidium myosin motor domain.";
 RL Biochemistry 36:11619-11628(1997).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE-98070605; PubMed-9405148;
 RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
 RT "X-ray crystal structure and solution fluorescence characterization
 of Mg₂(3')-O-(N-methylanthranilloyl) nucleotides bound to the
 Dictyostelium discoidium myosin motor domain.";
 RL J. Mol. Biol. 274:394-407(1997).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
 LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 (MLC-2).
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
 CORTEX.
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEMOROSIN (LM) AND 1 HEAVY MEMOROSIN (HM). IT CAN BE FURTHER
 SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- DOMAIN: THE ROD-LIKE TAIL SPOUNCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTF: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
 THE ACTIN-ACTIVATED ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
 ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
 POSITION (688).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M14628; AAA33227.1; -
 DR PIR: A26535; A26535.
 DR PIR: S00250; S00250.
 DR PDB: 1MMA: 03-DEC-97.
 DR PDB: 1HMD: 17-AUG-96.
 DR PDB: 1MMG: 03-DEC-97.
 DR PDB: 1MMN: 03-DEC-97.
 DR PDB: 1MNE: 17-AUG-96.
 DR PDB: 1VOM: 23-DEC-96.
 DR PDB: 1LVK: 28-JAN-98.
 DR Dictydb: DD01008; mbca.
 DR Interpro: IPR000048; IQ.
 DR Interpro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR Prodom: PD000355; myosin_head; 1.
 DR SMART: SM00242; MISC; 1.
 DR SMART: SM00015; IQ; 1.
 DR PROSITE: PS50096; IQ; 1.
 KW Myosin: Coiled coil; Actin-binding; ATP-binding; 3D-structure;
 KM Calmodulin-binding; Methylation; Alkylation; Phosphorylation;
 FT DOMAIN 1 761 MYOSIN HEAD-LIFE.
 FT DOMAIN 762 791 IQ.
 FT NP_BIND 817 2116 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP.
 FT DOMAIN 638 660 ACTIN-BINDING.
 FT DOMAIN 738 752 ACTIN-BINDING.
 FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).

FT MOD_RES 678 678 ALKYLATION (SH-1).
 FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MCK).
 FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MCK).
 FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MCK).
 SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;
 Query Match 8.6%; Score 159; DB 1; Length 2116;
 Best Local Similarity 21.2%; Pred. No. 3.6;
 Matches 98; Conservative 73; Mismatches 153; Indels 138; Gaps 18;
 QY 9 DGTMSANSDESVK-----PULTEISKITPSNVILAVKREVALLSIDELAAKIG 62
 DB 964 DGSDTISRLKIKDELQKVEELTFSESKDKGYL--EKTRVRLQSLDLYTR--- 1018
 QY 63 KKHQNGLDTEYNHNGSLAGAYASTLIKRLDKGNGLKEKIDAKKCEFTPTNKL 122
 DB 1019 -----LDSETKDSSELLRQKKLEELKQVQALAE-----TAAKLAQEAANKKL 1064
 QY 123 KEKHTDLGKEGVTDADAKKAILKTNGT-----KTGA----- 154
 DB 1065 QGEYTELNEKFNSVARSNVEKSKKLTLESQLVAVNNELDEEKKRNDALKEKKKALDAML 1124
 QY 155 EELGKLFESE-----VEVLSKAKEMLVANVETSPVY----- 187
 DB 1125 EEKKDLESTGEGKKSILYDLKVOESDMALRNOISELSTIKKEKIKTLEGEVARLQ 1184
 QY 188 ---AESPAMVNSGKD-----GNTSANSDESVKGPULTEISKITPSNVILAVKE 236
 DB 1185 GELAEQGLAKSNVEKQKKVELDLEKSNQALAEAKQALDKLKKLQGE-----LSE 1238
 QY 237 VETLLTSLDELAKAIGKKI-----NDVSLDNEDHNGSLISGAVYLSNLITK 284
 DB 1239 VQF-----QLSEANNKNVNSDSTNKHLETSFNNKLLEAEQAK-----QALEKK 1284
 QY 285 KI---SAIKDSGLKAEIKAKKCEFTAKLGEHTDGLKEGVTDNA--KKAILKTND 340
 DB 1285 RLGESEELKAVNQLEEKQKSNKREKRVYDLKEVSEL-KDQIEEVASKKAVTEAKK 1343
 QY 341 KTKGADELEKLF-----ESVKNLS--KAKEEMLTNSYKE 372
 DB 1344 KESELDLEIKRQYADVSSRDKSVQGLTKLQAKNEBELRNTAE 1385
 RESULT 13
 LMB1_MOUSE STANDARD; PRT; 1786 AA.
 AC P02469;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
 GN LAMB-1 OR LAMB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87147212; PubMed-3493487;
 RA Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;
 RT "Sequence of the cDNA encoding the laminin B1 chain reveals a
 multidomain protein containing cysteine-rich repeats.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).
 RN [2]
 RP SEQUENCE OF 1292-1786 FROM N.A.
 RX MEDLINE-85051302; PubMed-6209134;
 RA Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
 RT "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
 coiled-coil alpha-helix.";
 RL EMBO J. 3:2355-2362(1984).
 RN [3]
 RP SEQUENCE OF 165-172; 539-547 AND 712-719.

FT	DOMAIN	1398	1430	DOMAIN ALPHA.
FT	DOMAIN	1431	1486	DOMAIN I.
FT	DOMAIN	1216	1315	COILED COIL (POTENTIAL).
FT	DOMAIN	1368	1388	COILED COIL (POTENTIAL).
FT	DOMAIN	1448	1778	COILED COIL (POTENTIAL).
FT	DISULFID	271	280	BY SIMILARITY.
FT	DISULFID	273	298	BY SIMILARITY.
FT	DISULFID	300	309	BY SIMILARITY.
FT	DISULFID	312	332	BY SIMILARITY.
FT	DISULFID	335	344	BY SIMILARITY.
FT	DISULFID	337	362	BY SIMILARITY.
FT	DISULFID	365	374	BY SIMILARITY.
FT	DISULFID	377	395	BY SIMILARITY.
FT	DISULFID	398	411	BY SIMILARITY.
FT	DISULFID	400	426	BY SIMILARITY.
FT	DISULFID	428	437	BY SIMILARITY.
FT	DISULFID	440	455	BY SIMILARITY.
FT	DISULFID	458	472	BY SIMILARITY.
FT	DISULFID	460	479	BY SIMILARITY.
FT	DISULFID	481	490	BY SIMILARITY.
FT	DISULFID	493	507	BY SIMILARITY.
FT	DISULFID	773	785	BY SIMILARITY.
FT	DISULFID	775	792	BY SIMILARITY.
FT	DISULFID	794	803	BY SIMILARITY.
FT	DISULFID	806	818	BY SIMILARITY.
FT	DISULFID	821	833	BY SIMILARITY.
FT	DISULFID	823	840	BY SIMILARITY.
FT	DISULFID	842	851	BY SIMILARITY.
FT	DISULFID	854	864	BY SIMILARITY.
FT	DISULFID	867	876	BY SIMILARITY.
FT	DISULFID	869	883	BY SIMILARITY.
FT	DISULFID	886	895	BY SIMILARITY.
FT	DISULFID	898	914	BY SIMILARITY.
FT	DISULFID	917	933	BY SIMILARITY.
FT	DISULFID	919	944	BY SIMILARITY.
FT	DISULFID	946	955	BY SIMILARITY.
FT	DISULFID	958	973	BY SIMILARITY.
FT	DISULFID	976	990	BY SIMILARITY.
FT	DISULFID	978	997	BY SIMILARITY.
FT	DISULFID	1000	1009	BY SIMILARITY.
FT	DISULFID	1012	1025	BY SIMILARITY.
FT	DISULFID	1084	1096	BY SIMILARITY.
FT	DISULFID	1086	1103	BY SIMILARITY.
FT	DISULFID	1105	1114	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1132	1144	BY SIMILARITY.
FT	DISULFID	1134	1151	BY SIMILARITY.
FT	DISULFID	1153	1162	BY SIMILARITY.
FT	DISULFID	1165	1176	BY SIMILARITY.
FT	DISULFID	1179	1187	INTERCHAIN (PROBABLE).
FT	DISULFID	1182	1192	INTERCHAIN (PROBABLE).
FT	DISULFID	1185	1785	INTERCHAIN (PROBABLE).
FT	CARBOHYD	120	120	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	519	519	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	677	677	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1041	1041	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1195	1195	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1279	1279	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1336	1336	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1343	1343	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1487	1487	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1533	1533	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1542	1542	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1643	1643	N-LINKED (GLCNAC. . .)
FT	CONFLICT	1531	1534	SGNA -> MEMB (1N REF. 2).
FT	CONFLICT	1749	1749	D -> N (1N REF. 2).
SQ	SEQUENCE	1786	196504	MM; 846671B7BFA1A74 CRC64;

```

OY 3 CUNSGKDG-NTSANSADSVKGPNTLEI-----SKTITDSNAVLAVKEVEALLSI 53
DB 1406 CSECEGCGPNCRTDECEKCGPGCGGLVTVAHSAWQKAMDFDRDVLALAEVEQLSKW 1465
OY 54 DELAAKAIKKTHONNGLDTEYNHNSLAGAVAI STLKQKLDGKNEGKRIKIDAAK 113
DB 1466 SBAKVRADAKONADVL-----LKTNAKTKVD--K 1495
OY 114 CSEFTNKLKEKHTDLEKGEVTDADAKEA-----ILKNTGKTGAEIKLFEPS-----VE 165
DB 1496 SNEDELRLNIKQIRNFL--TEDSADLDISEAVANEVLSGNAST--POQLQNTPTDIEREP 1552
OY 166 VLSKAKEMLANSVKEELISPVVAESPAMVNSGKDGNTSANSADSVKGPNTLEIS--KK 223
DB 1553 TLSQVEVILQOASADIARAELEELAEKRAKSKSTDVKTADWYKFALEAEKQVAAEKA 1612
OY 224 ITESNAVVAKEVEETLLTSLD-----ELAKAIKKTKKKNVSLDNEADHNGSL 272
DB 1613 IKOADE--DIOGTOMLLTSLSEETAASEETLTMAOSQIRKERNVEELKRAKQON---- 1665
OY 273 SGAYLISNLITKKAISKDSE-----LKAETEKAKCSEEFKAKLGEHTDLEKGEVTD 327
DB 1666 SCE---AEYIEKVVSVKNADAVKKTLDGELEDEKRYKKEVSLIAOKTEESADARRKAE 1722
OY 328 DNAAKAIKLTNNDKTKGADELEKFE-SVKNLSKAKEK--LTNSVKEL 373
DB 1723 ONEAKTLLAQANSKIQLELDERKEDNOKYLEDKAQLVLEGEVRS 1771

```

RESULT 14
SCA4_RICAK STANDARD: PRT: 998 AA.

```

AC 09AIX9;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE ANTIGENIC HEAT-STABLE 120 KDA PROTEIN (PS120) (120 KDA ANTIGEN)
DE (PROTEIN PS 120) (FRAGMENT).
GN SCA4 OR D.
OS Rickettsia akari.
OC Bacteria: Proteobacteria: alpha subdivision: Rickettsiales:
OC Rickettsiaceae: Rickettsiidae: Rickettsia.
OX NCBI_TaxID=786;
RN [1]
RP SEQUENCE FROM N.A.
RA Sekeyova Z., Roux V., Raoult D.;
RT 'Phylogenetic analysis of Rickettsia spp. by comparing sequence of the
RT 'gene D' coding for an intracytoplasmic protein,"
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF213016; AAK30691.1; -
KM Antigen.
FT NON_TER 1
FT NON_TER 998
SQ SEQUENCE 998 AA: 109328 MW: FCEBA3AC62DE5BDS CRC64;

```

Query Match 8.5%; Score 157; DB 1; Length 998;
Best Local Similarity 21.1%; Pred. No. 1.9;
Matches 104; Conservative 65; Mismatches 139; Indels 184; Gaps 23;
OY 5 NSG-----KQGNSTANSADSV-----KGPNTLEISKIKIDSNVAVL--LAVKEVEAL 49
1:1 11111 : :1: 1 : :1: : :1: 1 1:1

```

DB 424 NAGLTREKDGNTQIDLINEATAILLNKEKQANFTILTKMVMNNNALPTDTKAVRNAV 483
OY 50 LSSIDEIAAKAIKKTHONNGLDTEYNHNSLAGAVAI STLKQKLDGKNEGK--KEK 107
DB 484 LETI-----KNQDTPDIEKSKMLEAVVAIT-----LNSENLTPRKOK 520
OY 108 IDAKKCSFTFKLKEKHTDLEKGEVTDADAKEA ILKTN-GRTKGAE--ELGKLFESV 164
DB 521 QOMLEKAVDVLSFKDDTSRAVAIDGITG---AVIKSNLSTKDKGTMLIAVGDKNVAS 575
OY 165 EYLSKAKEMLANSV-----KELTSP-----VVASPAVAVNSGKDG----- 201
DB 576 E-LSNAEKQDLGSLVKKEVETKILSPEDQQLMOONLDKTTAEQTKNDNTTEVOGLIAP 634
OY 202 --NTSANS-----DESYGPNLTSEIKITES----- 227
DB 635 AFNTIKTAIAIQVTKRVLDSPITAEIKGPTLSIFIVAESPLNVQDKTDIVKGMGEAI 694
OY 228 -----NAVVAKEVE-----LITS-----IDELAKAI 251
DB 695 ASHRTMAPTKKIAIISVETGVAKSITDLEDKLMTKGLVDGYEDKANBELTSEMMAV 754
OY 252 GKRIKNDVSLDNEADHNGSLISGAYLISNLITKKAISKDSEELKAEIKAKCSEBFTA 311
DB 755 SKGVDN-----STAIPEDKQALKRAAS-EAALDRA---TONFTE 769
OY 312 KLGEHTDLEKGEVTDNAAKA-----ILKTNNDTKGAE--LEKLFESV 356
DB 790 GLKGQNLDEPKR--RDYIYKKAODIAVALKNVYTVVDANPEKREVSEEVMMKTSILN 847
OY 357 NLSKAKEMLTN 368
DB 848 DISKIAIEKYN 859

```

RESULT 15

MYS_AEQIR ID MYS_AEQIR STANDARD: PRT: 1938 AA.

```

AC P24733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN HEAVY CHAIN, STRIATED MUSCLE.
OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Periomorpha; Pectinoidae;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-adductor muscle;
RC MEDLINE=92011595; PubMed=1917970;
RX Mylray L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Complete primary structure of a scallop striated muscle myosin heavy
RT chain. Sequence comparison with other heavy chains reveals regions
RT that might be critical for regulation."
RT J. Biol. Chem. 265:18469-18476(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Tissue-adductor muscle;
RC MEDLINE=91088319; PubMed=2263488;
RX Mylray L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
RT muscle myosin heavy chain."
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;
RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA Szent-Gyorgyi A.G., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2.8-A
RT resolution."
RL Nature 368:306-312(1994).
RN [4]

```

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
 RA MEDLINE-96419133; PubMed-8805510;
 RA Houdusse A., Cohen C.;
 RT "Structure of the regulatory domain of scallop myosin at 2-A
 RT resolution: implications for regulation";
 RL Structure 4:21-32(1996).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X55714; CAA39247.1;
 DR PIR; S13557; S13557.
 DR PIR; A40997; A40997.
 DR PDB; 1SCM; 30-APR-94.
 DR PDB; 1WDC; 11-JUL-96.
 DR InterPro; IPR000048; IQ.
 DR InterPro; IPR002928; Myosin_tail.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001609; myosin_head.
 DR Pfam; PF00612; IQ; 1.
 DR Pfam; PF00563; myosin_head; 1.
 DR PRINTS; PF01576; Myosin_tail; 1.
 DR PRINTS; P00193; MYOSINHEAVY.
 DR PRODOM; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PSS0096; IQ; 1.
 DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
 FT DOMAIN 1 777
 FT DOMAIN 778 805
 FT DOMAIN 836 1938
 FT NE_BIND 176 183
 FT MOD_RES 693 693
 FT MOD_RES 703 703
 FT HELIX 778 821
 FT TURN 822 823
 FT HELIX 825 833
 FT TURN 834 835
 SQ SEQUENCE 1938 AA; 222821 MW; A5CCE4127D1A4896 CRC64;
 Query Match 8.5%; Score 157; DB 1; Length 1938;
 Best Local Similarity 22.28; Pred. No. 4;
 Matches 87; Conservative 61; Mismatches 166; Indels 78; Gaps 13;
 QY 18 DESVKGPNLFEISKRTSDNSAVLAVKEVALLSSIDEIAAKGKIRHONGIDT-----73
 DB 925 DEEDAAADLEGIKKMEADNA-----NLKDDIDLENTLOKAEOAKHKNQDQISTLOGE 978
 QY 74 ---EYNHNGSLAGAYASTLIKOKLDGKNGELKEKIDAKKCEFTTNKLEKEHHDG 130
 DB 979 ISQODEHIGTKNKERKALEEANKKTSLSQAE--EDCKNLNKLKAKKLEQALDELEDNLE 1036
 QY 131 KEGYTDADAKFAIKTNGTGTGKAEEELGKLFESYEYLSKAKEMLANSVKELTSPVAES 190
 DB 1037 REKKVRGDVEKA-----KRKVEQDLKSTQENVEDLERVKRE-LEENVRKKEAEI-----1084

QY 191 PAMVNGSGKQONTSANSADESVK--GPNLFEISKRTI-TESNAVVLAVKEVETLLTSIDEL 247
 DB 1085 -SINSKLEDEQNLIVSOLQRRIRKELQARIELEEELEAEERNARAKVEKQRAELNELEEL 1143
 QY 248 AKAIK-----KIKNDV---SLDNEADHNGSLISGATVLSNLTG 284
 DB 1144 GERLDEAGATSAQIELNKKREAEELKIRPDLSEASLOHEAO-----ISALRRK 1192
 QY 285 KISAIKSGELKAEIEKAKKCEFTAKLKGHTDGGKGYVDNNAKKAILTNNDXTKG 344
 DB 1193 HODANEMADQVDLOKVKSKLEKDKDLKREMDLESG-----MTHMKKNG 1240
 QY 345 ADE-LKLFESYKNLSKAKEMLTNSVKELTS 375
 DB 1241 CSEKVKQFESQMSDLNARLEDSQRSINELQS 1272

Search completed: March 18, 2002, 10:11:42
 Job time: 960 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:32 ; Search time 55.5 Seconds
(without alignments)
153.266 Million cell updates/sec

Title: US-09-596-746A-28

Perfect score: 1859

Sequence: 1 MACNNSGKDGNTSANSNADSE.....KNLSKAKEMLTNSVKELTS 378

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212257 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	968.5	52.1	466	4	US-08-235-836C-110 Sequence 110, App
2	945	50.8	210	1	US-08-158-353-3 Sequence 3, Appl
3	938	50.5	210	4	US-08-209-603E-15 Sequence 15, Appl
4	933	50.2	209	4	US-09-196-293-15 Sequence 15, Appl
5	932	50.1	210	4	US-08-235-836C-30 Sequence 30, Appl
6	929	50.0	210	4	US-08-235-836C-122 Sequence 122, App
7	915	49.2	466	4	US-08-235-836C-107 Sequence 107, App
8	685	36.8	212	1	US-08-158-353-2 Sequence 2, Appl
9	670	36.0	212	1	US-08-031-295-2 Sequence 2, Appl
10	670	36.0	212	4	US-07-903-580-2 Sequence 2, Appl
11	631	33.9	212	1	US-08-158-353-4 Sequence 4, Appl
12	629	33.8	212	4	US-09-196-293-11 Sequence 11, Appl
13	629	33.8	212	4	US-08-209-603E-11 Sequence 11, Appl
14	626	33.7	212	4	US-08-235-836C-34 Sequence 34, Appl
15	614.5	33.1	209	4	US-08-235-836C-32 Sequence 32, Appl
16	592.5	31.9	207	4	US-08-235-836C-36 Sequence 36, Appl
17	171	9.2	194	4	US-09-364-083-2 Sequence 2, Appl
18	164	8.8	2482	1	US-08-328-254-6 Sequence 6, Appl
19	163	8.8	3248	1	US-08-353-700-1 Sequence 1, Appl
20	163	8.8	3248	5	PCT-US95-16216-1 Sequence 1, Appl
21	155	8.3	1388	2	US-08-685-576-1 Sequence 1, Appl
22	151.5	8.1	1388	2	US-08-685-576-4 Sequence 4, Appl
23	148	8.0	3111	2	US-08-460-309-4 Sequence 4, Appl
24	148	8.0	3111	2	US-08-125-077-4 Sequence 4, Appl
25	146	7.9	688	3	US-09-141-047-8 Sequence 8, Appl
26	145.5	7.8	1196	1	US-08-144-121-4 Sequence 4, Appl
27	145.5	7.8	1196	2	US-08-735-893-4 Sequence 4, Appl

ALIGNMENTS

28	144.5	7.8	1164	4	US-08-923-992A-2	Sequence 2, Appl
29	144	7.7	1713	3	US-08-600-982-24	Sequence 24, Appl
30	144	7.7	1713	5	PCT-US94-10261A-24	Sequence 24, Appl
31	143	7.7	630	4	US-08-973-462-9	Sequence 9, Appl
32	142.5	7.7	1164	4	US-08-923-992A-10	Sequence 10, Appl
33	142	7.6	1507	3	US-08-929-329-5	Sequence 5, Appl
34	142	7.6	1786	4	US-08-973-462-8	Sequence 8, Appl
35	139.5	7.5	1588	5	PCT-US93-07261-11	Sequence 11, Appl
36	139.5	7.5	1653	5	PCT-US93-07261-16	Sequence 16, Appl
37	138	7.4	1098	4	US-08-923-992A-8	Sequence 8, Appl
38	137	7.4	641	4	US-08-961-083-160	Sequence 160, App
39	136.5	7.3	1038	4	US-09-541-782-4	Sequence 4, Appl
40	135.5	7.3	1104	4	US-08-923-992A-6	Sequence 6, Appl
41	135.5	7.3	1128	4	US-08-923-992A-6	Sequence 6, Appl
42	135	7.3	573	4	US-08-235-836C-112	Sequence 112, App
43	134	7.2	376	6	5180810-1	Sequence 23, Appl
44	132	7.1	1561	3	US-08-894-017-23	Sequence 23, Appl
45	131.5	7.1	695	1	US-08-127-499A-23	Sequence 23, Appl

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110
Query Match 52.1%; Score 968.5; DB 4; Length 466;
Best Local Similarity 59.5%; Pred. No. 5.2e-62;
Matches 229; Conservative 32; Mismatches 93; Indels 31; Gaps 9;
OY 1 MACNNSGKDGNTSANSNADSEVKGPNLFEINKKIDSNVLLAVKEVALLSIDIAKA 60

Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLLAVEALLSIDETIAKA 76
QY 61 IGGKIHONNGLDTEENNHNHSLAGAVAI STLKOKLDGLKNGEKIDAKKCSSETFN 120
Db 77 IGGKIHONNGLDTEENNHNHSLAGAVAI STLKOKLDGLKNGEKIDAKKCSSETFN 120
QY 121 KLRKHTDLGREGVTDADAKAEIILKANGTKRGAELGKLFESVYLSKAEMLANSVK 180
Db 137 KLRKHTDLGREGVTDADAKAEIILKANGTKRGAELGKLFESVYLSKAEMLANSVK 180
QY 181 ELTSPVVAESPKKPVNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLLAVEALLSIDETIAKA 76
Db 197 ELTSPVVAESPKKPVNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLLAVEALLSIDETIAKA 76
QY 239 EVETILTS-----IDLAKAIGKRTKNDVSLDNEADHNSISAVYI--SNLITRK 289
Db 253 KLEKGTSDKNGSGVLEGVAKSKVKILTS-----DLGOTTLVEKEDGKTLVSKV 307
QY 290 SAIDSGELKAEIKRACSEFTAKLGEHTDLGREGVTDNNAKAILTNNDKTGAD 349
Db 308 TS-KDKSSTEKEFNEKEVESEKILTRADG--TRLEYTGIKSDSGKAEVILKGVILEGL 364
QY 350 ELEKLFESV-----KNLSKA 366
Db 365 TAEXTLVYEGVTLSKNISKSGE 389

RESULT 2
US-08-158-353-3
Sequence 3, Application US/08158353
Patent No. 3620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
TELECOMMUNICATION INFORMATION: DCT93-05
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9340
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-3

Query Match 50.8%; Score 945; DB 1; Length 210;
Best Local Similarity 97.9%; Pred. No. 8.9e-61;
Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNPLTEINKRTIDSNVAVLLAVEALLSIDETIAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLLAVEALLSIDETIAKA 76
QY 61 IGGKIHONNGLDTEENNHNHSLAGAVAI STLKOKLDGLKNGEKIDAKKCSSETFN 120
Db 77 IGGKIHONNGLDTEENNHNHSLAGAVAI STLKOKLDGLKNGEKIDAKKCSSETFN 120
QY 121 KLRKHTDLGREGVTDADAKAEIILKANGTKRGAELGKLFESVYLSKAEMLANSVK 180
Db 137 KLRKHTDLGREGVTDADAKAEIILKANGTKRGAELGKLFESVYLSKAEMLANSVK 180
QY 181 ELTSPVVAESPKKPVNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLLAVEALLSIDETIAKA 76
Db 197 ELTSPVVAESPKKPVNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVAVLLAVEALLSIDETIAKA 76

RESULT 3
US-08-209-603E-15
Sequence 15, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: RUCHS, RENATE
APPLICANT: WILKE, BETTINA
APPLICANT: PREAC-MURISIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAEFFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
CLASSIFICATION: 436
PRIOR APPLICATION: 436
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 697-3355
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE: N/A
ORGANISM: B. BURGDOFFERT
IMMEDIATE SOURCE:
LIBRARY: DSM 5662

POSITION IN GENOME: N/A
FEATURE: N/A
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

Query Match 50.5%; Score 938; DB 4; Length 210;
Best Local Similarity 97.4%; Pred. No. 2.8e-60;
Matches 189; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSSIDEIAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAKA 76
OY 61 IGKRIHONNGLDTENNHNSGLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
DB 77 IGKRIHONNGLDTENNHNSGLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 136
OY 121 KLEKHTDGLKEGYTDADAKAELIKANGTKTKGAELGKLFESVEVLSKAKEMLANSVK 180
137 KLEKHTDGLKEGYTDADAKAELIKANGTKTKGAELGKLFESVEVLSKAKEMLANSVK 196
OY 181 ELTSPVVAESPKP 194
DB 197 ELTSPVVAESPKP 210

RESULT 4
US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutacheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Praec-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293
CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FASTED for Windows Version 4.0
SEQ ID NO 15
LENGTH: 209
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match 50.2%; Score 933; DB 4; Length 209;
Best Local Similarity 97.9%; Pred. No. 6.4e-60;
Matches 188; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSSIDEIAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAKA 76
OY 61 IGKRIHONNGLDTENNHNSGLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
DB 77 IGKRIHONNGLDTENNHNSGLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 136

OY 121 KLEKHTDGLKEGYTDADAKAELIKANGTKTKGAELGKLFESVEVLSKAKEMLANSVK 180
DB 137 KLEKHTDGLKEGYTDADAKAELIKANGTKTKGAELGKLFESVEVLSKAKEMLANSVK 196
OY 181 ELTSPVVAESPK 192
DB 197 ELTSPVVAESPK 208

RESULT 5
US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-Apr-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-30

Query Match 50.1%; Score 932; DB 4; Length 210;
Best Local Similarity 96.9%; Pred. No. 7.6e-60;
Matches 188; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSSIDEIAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAKA 76
OY 61 IGKRIHONNGLDTENNHNSGLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 120
DB 77 IGKRIHONNGLDTENNHNSGLAGAYAISTLIKOKLDGKNEGLKEKIDAKKCEFTFN 136
OY 121 KLEKHTDGLKEGYTDADAKAELIKANGTKTKGAELGKLFESVEVLSKAKEMLANSVK 180
DB 137 KLEKHTDGLKEGYTDADAKAELIKANGTKTKGAELGKLFESVEVLSKAKEMLANSVK 196
OY 181 ELTSPVVAESPKP 194

Db 197 ELTSPVVAESPCKP 210

RESULT 6

US-08-235-836C-122

Sequence 122, Application US/08235836C

Patent No. 6248562

GENERAL INFORMATION:

APPLICANT: Dunn, John J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/235,836C

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-7338

INFORMATION FOR SEQ ID NO: 122:

SEQUENCE CHARACTERISTICS:

LENGTH: 588 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-122

Query Match 50.0%; Score 929; DB 4; Length 588;

Best Local Similarity 89.7%; Pred. No. 4.7e-59;

Matches 191; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Db 4 NNSGKDGNTSANSADSEYKGPMLTEINKITDSNAVLLAVEVALLSSIDEIAKAIGK 63

Db 286 NNSGKDGNTSANSADSEYKGPMLTEISKITDSNAVLLAVEVALLSSIDEIAKAIGK 345

Db 64 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGKEKIDAAKCSFTTNKIK 123

Db 346 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGKEKIDAAKCSFTTNKIK 405

Db 124 EKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 183

Db 406 EKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 465

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

Db 466 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 498

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

Db 466 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 498

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

Db 466 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 498

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

Db 466 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 498

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

Db 466 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 498

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

Db 466 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 498

Db 184 SPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKPSPVVAESPCKP 216

GENERAL INFORMATION:

APPLICANT: Dunn, John J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brookhaven National Laboratory

STREET:

CITY: Upton

STATE: NY

COUNTRY: USA

ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/235,836C

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324

REFERENCE/DOCKET NUMBER: BNL93-28A

TELEPHONE: (516) 282-7338

TELEFAX: (516) 282-7338

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 466 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-235-836C-107

Query Match 49.2%; Score 915; DB 4; Length 466;

Best Local Similarity 97.4%; Pred. No. 3.5e-58;

Matches 186; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 4 NNSGKDGNTSANSADSEYKGPMLTEINKITDSNAVLLAVEVALLSSIDEIAKAIGK 63

Db 276 NNSGKDGNTSANSADSEYKGPMLTEISKITDSNAVLLAVEVALLSSIDEIAKAIGK 335

Db 64 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGKEKIDAAKCSFTTNKIK 123

Db 336 KIHONNGIDTENNHNGSLAGAYASTLIKOKLDKNEGKEKIDAAKCSFTTNKIK 395

Db 124 EKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 183

Db 396 AKHTDLGEGYTDADAKAAILIKANGTKTGAEEGLKLFESVEVLSKAKEMLANSVELT 455

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

Db 184 SPVVAESPCKP 194

Db 456 SPVVAESPCKP 466

STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-2

Query Match 36.8%; Score 685; DB 1; Length 212;
Best Local Similarity 74.1%; Pred. No. 3.8e-42;
Matches 146; Conservative 19; Mismatches 28; Indels 4; Gaps 3;

OY 1 MACNSGKDGNTSANSADSVKGNLTKETKITSNAVLAVKVEYALLSSIDEIAAK 60
DB 17 ISCNNSGKDGNTSANSADSVKGNLTKETKITSNAVLAVKVEYALLSSIDEIAAK 76
OY 61 IGKKIHONNGLDTEENNHNHSLAGAVYSTLIKOKLDGLKN-EGLEKIDAAKCSSEPT 119
DB 77 IGNLIAQ-NGLNAGANONGSLAGAVYSTLIAEKLDGLKNEBELKEKIEDAKKCKKAT 135
OY 120 NKLAKKHDTLQ--KEGVTDADAKAAILKANGTKTGAEELGLFESVEVLSAKAKEMLAN 177
DB 136 DKLKSSHAELGANGAATDANAKAAILKTNGTKDGAQELERLFESVKNLSKAAOETLNN 195
OY 178 SVKELTSPVVAESPCKP 194
DB 196 SVKELTSPVVAENPKP 212

RESULT 9
US-08-031-295-2
Sequence 2, Application US/08031295
Patent No. 5530103
GENERAL INFORMATION:
APPLICANT: LIVERY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,295
FILING DATE: 19930312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903,580
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/142 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-031-295-2

Query Match 36.0%; Score 670; DB 1; Length 212;
Best Local Similarity 73.0%; Pred. No. 4.5e-41;
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

OY 1 MACNSGKDGNT-SANSADSVKGNLTKETKITSNAVLAVKVEYALLSSIDEIAK 59
DB 17 ISCNNSGKGDSDS-STNPADESAGKPNLTKETSKITSNAVLAVKVEYALLSSIDEIAK 76
OY 60 AIGKKIHONNGLDTEENNHNHSLAGAVYSTLIKOKLDGLKNG-LKEKIDAAKCSSEPT 118
DB 77 AIGKKIHONNGLDANNDKNSLAGAVYSTLIKOKLDGLKNGELKAKIEDAKKCSSEPT 136
OY 119 TNKLAKKHDTLQEGVTDADAKAAILKANGTKTGAEELGLFESVEVLSAKAKEMLAN 178
DB 137 TKRLAGHQLGIDGATDNDKSKAAILKTNGTKTGAEELGLFESVLSKAAOEAASANS 196
OY 179 VKELTSPVVAESPCKP 194
DB 197 VKELTSPVVAETPKP 212

RESULT 10
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
GENERAL INFORMATION:
APPLICANT: LIVERY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 19920625
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/131 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-580-2

Query Match 36.0%; Score 670; DB 4; Length 212;
Best Local Similarity 73.0%; Pred. No. 4,56-41;
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;
QY 1 MACNSGKGN-SANSADSVKGNLTKITDSDNAVLAKVEYALLSIDETIAK 59
DB 17 ISCNNSGKGDSTNPADSAKGNLTKITDSDNAVLAKVEYALLSIDETIAK 59
QY 60 AIGKKIHONNGIDENNNHNSLAGAVAI STLKOKIDGLKN-EGLEKIDAKKCESEF 118
DB 77 AIGKKIQONNGIGANADNSLGAVAISTLTETKALKKNGELKAKIEDAKKCESEF 136
QY 119 TNKLEKHTDCKEGVDADAKRAILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANS 178
DB 137 TKKLAGHAQIDGATDNDKSKAILKTNGTKTKGAEEIGKLFESVEVLSKAKEMLANS 178
QY 179 VKELTSPVAESPKRP 194
DB 197 VKELTSPVAETPKRP 212
RESULT 11
US-08-158-353-4
Sequence 4, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
STREET: Hamilton, Brook, Smith & Reynolds, P.C.
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-6240
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match 33.8%; Score 631; DB 1; Length 212;
Best Local Similarity 69.9%; Pred. No. 2,8e-38;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;
QY 1 MACNSGKGN-SANSADSVKGNLTKITDSDNAVLAKVEYALLSIDETIAK 59
DB 17 ISCNNSGKGDSTNPADSAKGNLTKITDSDNAVLAKVEYALLSIDETIAK 59
QY 60 AIGKKIHONNGIDENNNHNSLAGAVAI STLKOKIDGLKN-EGLEKIDAKKCESEF 118
DB 77 AIGKKIQONNGIGANADNSLGAVAISTLTETKALKKNGELKAKIEDAKKCESEF 136
QY 119 TNKLEKHTDCKEGVDADAKRAILKANGTKTKGAEEIGKLFESVEVLSKAKEMLANS 178
DB 137 TKKLAGHAQIDGATDNDKSKAILKTNGTKTKGAEEIGKLFESVEVLSKAKEMLANS 178
QY 179 VKELTSPVAESPKRP 194
DB 197 VKELTSPVAETPKRP 212

RESULT 12
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: 738,001US2
CURRENT APPLICATION NUMBER: US/09/196,293
EARLIER APPLICATION NUMBER: US-08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE B39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match 33.8%; Score 629; DB 4; Length 212;

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-34

Query Match 33.7%; Score 626; DB 4; Length 212;
Best Local Similarity 69.4%; Pred. No. 6,3e-38;
Matches 136; Conservative 16; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNSGKDGNT-SANSDESVKGNLLEINKKTTDSNAVLAVKEVEALLSIDETIAK 59
DB 17 ISCSNSG--GDTASTNPDESAGPNLTVISKRTTDSNAVLAVKEVEALLSIDETIAK 76
QY 60 AIGKTIHONNGIDTENNHNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFT 118
DB 77 AIGKTIHONNGIDTENNHNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFT 136
DB 119 TNKLEKHTDLGEGVTDAADAKKAIKANGTKTKGAELGKLFESVEVLSKAKEMLANS 178
137 TNKLSGSHADLGKQDATTDDHAKAAILKTHATTDKGAKEFKLFESVEVLSKAQVALTNS 196
QY 179 VKELTSPVAESPKP 194
DB 197 VKELTSPVAESPKP 212

RESULT 15

US-08-235-836C-32
Sequence 32, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Peptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-32

Query Match 33.1%; Score 614.5; DB 4; Length 209;
Best Local Similarity 68.2%; Pred. No. 4.1e-37;

Matches 133; Conservative 20; Mismatches 39; Indels 3; Gaps 2;
QY 1 MACNSGKDGNTSANSADSVKGNLLEINKKTTDSNAVLAVKEVEALLSIDETIAK 60
DB 17 ISCSNSG--GDTASTNPDESAGPNLTVISKRTTDSNAVLAVKEVEALLSIDETIAK 74
QY 61 IGRVTHONNGIDTENNHNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFT 119
DB 75 IGRVTHONNGIDTENNHNHNSLLAGAVASTLIKOKLDGLKN-EGLEKIDAAKCSFT 134
QY 120 NKLKHTDLGEGVTDAADAKKAIKANGTKTKGAELGKLFESVEVLSKAKEMLANSV 179
DB 135 NKLKSHADLGVAATAATDDHAKAAILKSNPTKDKGAKALKDLSSEVSEVLSKAQVALTNSV 194
QY 180 KELTSPVAESPKP 194
DB 195 KELTSPVAESPKP 209

Search completed: March 18, 2002, 09:55:32
Job time: 345 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:42 : Search time 621.2 Seconds
(without alignments)
168.954 Million cell updates/sec

Title: US-09-596-746A-28

Perfect score: 1859
Sequence: 1 MACNNSGKDGTSANSADES.....KNLSAKAKELTNVKEKLTLS 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06.COMB.pep:*

2: /cgn2_6/ptodata/2/paa/US07.COMB.pep:*

3: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*

4: /cgn2_6/ptodata/2/paa/US081.COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US082.COMB.pep:*

6: /cgn2_6/ptodata/2/paa/US083.COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US084.COMB.pep:*

8: /cgn2_6/ptodata/2/paa/US085.COMB.pep:*

9: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*

10: /cgn2_6/ptodata/2/paa/US087.COMB.pep:*

11: /cgn2_6/ptodata/2/paa/US088.COMB.pep:*

12: /cgn2_6/ptodata/2/paa/US089.COMB.pep:*

13: /cgn2_6/ptodata/2/paa/US090.COMB.pep:*

14: /cgn2_6/ptodata/2/paa/US091.COMB.pep:*

15: /cgn2_6/ptodata/2/paa/US092.COMB.pep:*

16: /cgn2_6/ptodata/2/paa/US093.COMB.pep:*

17: /cgn2_6/ptodata/2/paa/US094.COMB.pep:*

18: /cgn2_6/ptodata/2/paa/US095.COMB.pep:*

19: /cgn2_6/ptodata/2/paa/US096.COMB.pep:*

20: /cgn2_6/ptodata/2/paa/US097.COMB.pep:*

21: /cgn2_6/ptodata/2/paa/US098.COMB.pep:*

22: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*

23: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*

24: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1859	100.0	378	19	US-09-596-746A-28
2	1859	100.0	378	19	US-09-596-746A-28
3	1834	98.7	401	19	US-09-596-746A-80
4	1834	98.7	402	19	US-09-596-746A-80
5	1810.5	97.4	375	19	US-09-596-746A-24
6	1805.5	97.1	374	19	US-09-596-746A-24
7	1802.5	97.0	398	19	US-09-596-746A-60
8	1802.5	97.0	399	19	US-09-596-746A-60
9	1614.5	86.8	377	19	US-09-596-746A-30

10	1614.5	86.8	377	19	US-09-596-746A-30	Sequence 30, Appl
11	1589.5	85.5	400	19	US-09-596-746A-82	Sequence 82, Appl
12	1589.5	85.5	401	19	US-09-596-746A-82	Sequence 82, Appl
13	1566	84.2	374	19	US-09-596-746A-26	Sequence 26, Appl
14	1561	84.0	373	19	US-09-596-746A-26	Sequence 26, Appl
15	1558	83.8	397	19	US-09-596-746A-62	Sequence 62, Appl
16	1558	83.8	398	19	US-09-596-746A-62	Sequence 62, Appl
17	1533.5	82.5	371	19	US-09-596-746A-34	Sequence 34, Appl
18	1528.5	82.2	370	19	US-09-596-746A-34	Sequence 34, Appl
19	1525.5	82.1	354	19	US-09-596-746A-64	Sequence 64, Appl
20	1525.5	82.1	355	19	US-09-596-746A-64	Sequence 64, Appl
21	1501.5	80.8	385	19	US-09-596-746A-56	Sequence 56, Appl
22	1496.5	80.5	384	19	US-09-596-746A-56	Sequence 56, Appl
23	1494.5	80.4	369	19	US-09-596-746A-40	Sequence 40, Appl
24	1493.5	80.3	408	19	US-09-596-746A-78	Sequence 78, Appl
25	1493.5	80.3	409	19	US-09-596-746A-78	Sequence 78, Appl
26	1489.5	80.1	368	19	US-09-596-746A-40	Sequence 40, Appl
27	1486.5	80.0	392	19	US-09-596-746A-72	Sequence 72, Appl
28	1486.5	80.0	393	19	US-09-596-746A-72	Sequence 72, Appl
29	1460.5	78.6	560	19	US-09-596-746A-52	Sequence 52, Appl
30	1460.5	78.6	560	19	US-09-596-746A-52	Sequence 52, Appl
31	1289	69.3	370	19	US-09-596-746A-36	Sequence 36, Appl
32	1284	69.1	369	19	US-09-596-746A-36	Sequence 36, Appl
33	1281	68.9	393	19	US-09-596-746A-66	Sequence 66, Appl
34	1281	68.9	394	19	US-09-596-746A-66	Sequence 66, Appl
35	1250	67.2	368	19	US-09-596-746A-42	Sequence 42, Appl
36	1245	67.0	367	19	US-09-596-746A-42	Sequence 42, Appl
37	1242	66.8	331	19	US-09-596-746A-74	Sequence 74, Appl
38	1242	66.8	332	19	US-09-596-746A-74	Sequence 74, Appl
39	1217	65.5	370	19	US-09-596-746A-32	Sequence 32, Appl
40	1213	65.3	378	19	US-09-596-746A-54	Sequence 54, Appl
41	1213	65.3	378	19	US-09-596-746A-54	Sequence 54, Appl
42	1212	65.2	369	19	US-09-596-746A-32	Sequence 32, Appl
43	1209	65.0	393	19	US-09-596-746A-68	Sequence 68, Appl
44	1209	65.0	394	19	US-09-596-746A-68	Sequence 68, Appl
45	1205	64.8	401	19	US-09-596-746A-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1

US-09-596-746-28

Sequence 28, Application US/09596746

GENERAL INFORMATION:

APPLICANT: Dattwyler, Raymond J.

APPLICANT: Selinost, Gerald

APPLICANT: Dykhuisen, Daniel

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and Borrelia afzelii That Cause Lyme Disease In Humans

FILE REFERENCE: 2631.1002-001

CURRENT APPLICATION NUMBER: US/09/596,746

CURRENT FILING DATE: 2000-06-16

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 28

LENGTH: 378

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Ospa Chimera

US-09-596-746-28

Query Match 100.0%; Score 1859; DB 19; Length 378;
Best Local Similarity 100.0%; Pred. No. 4, 5e-123;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTSANSADESKGNLTIEINKKITDSNAVLAVKEVALLSIDEIAKA 60

```

Db      1  MACNNSGKDGNTSANSADSVKGNLTETINKTIDTSNAVLAVKEVEALLSSIDEIAAK
Qy      61  IGKRIHONNGLDTEENHNGSLLAGAYASTILIKOKLDGKNGLEKIDAKKCSSEFTN
Db      61  IGKRIHONNGLDTEENHNGSLLAGAYASTILIKOKLDGKNGLEKIDAKKCSSEFTN
Qy      121  KLEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKFESEVLSKAKEMLANSVK
Db      121  KLEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKFESEVLSKAKEMLANSVK
Qy      181  ELTSPVAESPCKPSVNNNGKDGNTSANSADSVKGNLTETISKITESNAVYLAVEV
Db      181  ELTSPVAESPCKPSVNNNGKDGNTSANSADSVKGNLTETISKITESNAVYLAVEV
Qy      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISATIDSGELKA
Db      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISATIDSGELKA
Qy      301  EIEKAKKCSSEFTAKLKGHTDGLKGEVTDADAKKAILKTNDKTGAELEKLFESVKN
Db      301  EIEKAKKCSSEFTAKLKGHTDGLKGEVTDADAKKAILKTNDKTGAELEKLFESVKN
Qy      361  LSKAKEMLTNSVKELTS 378
Db      361  LSKAKEMLTNSVKELTS 378

```

RESULT 2

```

US-09-596-746a-28
/ Sequence 28, Application US/09596746a
/ GENERAL INFORMATION:
/ APPLICANT: Dattwyler, Raymond J.
/ APPLICANT: Selmost, Gerald
/ APPLICANT: Dykhuzen, Daniel
/ APPLICANT: Luft, Benjamin J.
/ TITLE OF INVENTION: Groups of Borrelia burgdorferi
/ FILE REFERENCE: 2631.1002-001
/ CURRENT APPLICATION NUMBER: US/09/596,746a
/ PRIOR FILING DATE: 2000-06-19
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ LENGTH: 378
/ TYPE: PRF
/ ORGANISM: Artificial Sequence
/ FEATURE: OTHER INFORMATION: Ospec Chimera
US-09-596-746a-28

```

```

Query Match
Best Local Similarity 100.0%; Score 1859; DB 19; Length 378;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  MACNNSGKDGNTSANSADSVKGNLTETINKTIDTSNAVLAVKEVEALLSSIDEIAAK
Db      1  MACNNSGKDGNTSANSADSVKGNLTETINKTIDTSNAVLAVKEVEALLSSIDEIAAK
Qy      61  IGKRIHONNGLDTEENHNGSLLAGAYASTILIKOKLDGKNGLEKIDAKKCSSEFTN
Db      61  IGKRIHONNGLDTEENHNGSLLAGAYASTILIKOKLDGKNGLEKIDAKKCSSEFTN
Qy      121  KLEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKFESEVLSKAKEMLANSVK
Db      121  KLEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKFESEVLSKAKEMLANSVK
Qy      181  ELTSPVAESPCKPSVNNNGKDGNTSANSADSVKGNLTETISKITESNAVYLAVEV
Db      181  ELTSPVAESPCKPSVNNNGKDGNTSANSADSVKGNLTETISKITESNAVYLAVEV

```

```

Db      181  ELTSPVAESPCKPSVNNNGKDGNTSANSADSVKGNLTETISKITESNAVYLAVEV
Qy      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISATIDSGELKA
Db      241  ETLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISATIDSGELKA
Qy      301  EIEKAKKCSSEFTAKLKGHTDGLKGEVTDADAKKAILKTNDKTGAELEKLFESVKN
Db      301  EIEKAKKCSSEFTAKLKGHTDGLKGEVTDADAKKAILKTNDKTGAELEKLFESVKN
Qy      361  LSKAKEMLTNSVKELTS 378
Db      361  LSKAKEMLTNSVKELTS 378

```

RESULT 3

```

US-09-596-746-80
/ Sequence 80, Application US/09596746
/ GENERAL INFORMATION:
/ APPLICANT: Dattwyler, Raymond J.
/ APPLICANT: Selmost, Gerald
/ APPLICANT: Dykhuzen, Daniel
/ APPLICANT: Luft, Benjamin J.
/ TITLE OF INVENTION: Groups of Borrelia burgdorferi
/ FILE REFERENCE: 2631.1002-001
/ CURRENT APPLICATION NUMBER: US/09/596,746
/ PRIOR FILING DATE: 2000-06-19
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ LENGTH: 401
/ TYPE: PRF
/ ORGANISM: Ospec Chimera
US-09-596-746-80

```

```

Query Match
Best Local Similarity 98.7%; Score 1834; DB 19; Length 401;
Matches 373; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy      2  ACNNSGKDGNTSANSADSVKGNLTETINKTIDTSNAVLAVKEVEALLSSIDEIAAK
Db      25  SCNNSGKDGNTSANSADSVKGNLTETINKTIDTSNAVLAVKEVEALLSSIDEIAAK
Qy      62  GKRTHONNGLDTEENHNGSLLAGAYASTILIKOKLDGKNGLEKIDAKKCSSEFTN
Db      85  GKRTHONNGLDTEENHNGSLLAGAYASTILIKOKLDGKNGLEKIDAKKCSSEFTN
Qy      122  LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKFESEVLSKAKEMLANSVK
Db      145  LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKFESEVLSKAKEMLANSVK
Qy      242  TLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISATIDSGELKA
Db      265  TLTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISATIDSGELKA
Qy      302  EIEKAKKCSSEFTAKLKGHTDGLKGEVTDADAKKAILKTNDKTGAELEKLFESVKN
Db      325  EIEKAKKCSSEFTAKLKGHTDGLKGEVTDADAKKAILKTNDKTGAELEKLFESVKN
Qy      362  SKAKEMLTNSVKELTS 378
Db      385  SKAKEMLTNSVKELTS 401

```

```
RESULT 4
US-09-596-746a-80
; Sequence 80, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 80
; LENGTH: 402
; TYPE: PRT
; ORGANISM: ospc Chimera
; 9-596-746a-80

Query Match          98.7%; Score 1834; DB 19; Length 402;
Best Local Similarity 98.9%; Pred. No. 2.9e-121;
Matches 373; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 26 SCNNSKDGNTSANSADSVKGNLTETISKTTDSNAVLAVKEVEALLSSIDEIAAKAI 85
OY 62 GKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCCSEFTFNK 121
DB 86 GKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCCSEFTFNK 145
OY 122 LKEKHTDLGEGVTDADAKAELIKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 146 LKEKHTDLGEGVTDADAKAELIKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVKE 205
OY 182 LSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAKEVE 241
DB 206 LSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAKEVE 265
OY 242 TLTTSIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTTRKISAIKDSGELKAE 301
DB 266 TLTTSIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTTRKISAIKDSGELKAE 325
OY 302 IEKAKKCESEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKNL 361
DB 326 IEKAKKCESEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKNL 385
OY 362 SKRAKEMLTNSVKELTIS 378
DB 386 SKRAKEMLTNSVKELTIS 402

RESULT 5
US-09-596-746a-24
; Sequence 24, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; OTHER APPLICATION NUMBER: US 60/140,042
```

```
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Ospc Chimera
; US-09-596-746a-24

Query Match          97.4%; Score 1810.5; DB 19; Length 375;
Best Local Similarity 98.1%; Pred. No. 1.2e-119;
Matches 371; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSSIDEIAAKA 60
DB 1 MACNNSGKDGNTSANSADSVKGNLTETISKTTDSNAVLAVKEVEALLSSIDEIAAKA 60
OY 61 IGKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCCSEFTFN 120
DB 61 IGKRIHONNGLDTEYNNHNSGLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCCSEFTFN 120
OY 121 KLEKHTDLGEGVTDADAKAELIKANGTKTKGAEELGKLFESVEVLSKAKEMLANSV 180
DB 121 KLEKHTDLGEGVTDADAKAELIKANGTKTKGAEELGKLFESVEVLSKAKEMLANSV 180
OY 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAKEVE 240
DB 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKTTESNAVLAKEVE 237
OY 241 ETLTSLIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTTRKISAIKDSGELKA 300
DB 238 ETLTSLIDELAKAIGKKIKNDVSLDNEADHNSGLISGAYLISNLTTRKISAIKDSGELKA 297
OY 301 EIEKAKKCESEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKN 360
DB 298 EIEKAKKCESEFTAKLKGHTDLGEGVTDNNAKAILKTNNDKTKGADELEKLFESVKN 357
OY 361 LSKRAKEMLTNSVKELTIS 378
DB 358 LSKRAKEMLTNSVKELTIS 375

RESULT 6
US-09-596-746-24
; Sequence 24, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Ospc Chimera
; US-09-596-746-24

Query Match          97.1%; Score 1805.5; DB 19; Length 374;
```

Best Local Similarity 98.1%; Pred. No. 2.7e-119;
Matches 370; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKTTSDNAVLLAVEVEALLSSIDEIAAKAI 61
Db 1 ACNNSGKDGNTSANSADSVKGNLTETINKKTTSDNAVLLAVEVEALLSSIDEIAAKAI 60
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLTKOKLDGKNEGKEKIDAAKCSSEFTFNK 121
Db 61 GKRIHONNGLDTEYNNHNSLAGAVAI STLTKOKLDGKNEGKEKIDAAKCSSEFTFNK 120
QY 122 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 181
Db 121 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 180
QY 182 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNLTETINKKTTESNAVYLAKEVE 241
Db 181 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGNLTETINKKTTESNAVYLAKEVE 237
QY 242 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAVILSNLTTRKISAIKDSGELKAE 301
Db 238 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAVILSNLTTRKISAIKDSGELKAE 297
QY 302 IEKAKKSEEFATAKLGHEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 361
Db 298 IEKAKKSEEFATAKLGHEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 357
QY 362 SKAKEMLTNSVKELTS 378
Db 358 SKAKEMLTNSVKELTS 374
```

RESULT 7

```
US-09-596-746-60
; Sequence 60, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Datwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chlamera
US-09-596-746-60
```

Query Match 97.0%; Score 1802.5; DB 19; Length 398;
Best Local Similarity 97.9%; Pred. No. 4.8e-119;
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKTTSDNAVLLAVEVEALLSSIDEIAAKAI 61
Db 25 SCNNSGKDGNTSANSADSVKGNLTETINKKTTSDNAVLLAVEVEALLSSIDEIAAKAI 84
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLTKOKLDGKNEGKEKIDAAKCSSEFTFNK 121
Db 85 GKRIHONNGLDTEYNNHNSLAGAVAI STLTKOKLDGKNEGKEKIDAAKCSSEFTFNK 144
QY 122 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 181
Db 145 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 204
```

```
QY 182 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNLTETINKKTTESNAVYLAKEVE 241
Db 205 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGNLTETINKKTTESNAVYLAKEVE 261
QY 242 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAVILSNLTTRKISAIKDSGELKAE 301
Db 262 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAVILSNLTTRKISAIKDSGELKAE 321
QY 302 IEKAKKSEEFATAKLGHEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 361
Db 322 IEKAKKSEEFATAKLGHEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 381
QY 362 SKAKEMLTNSVKELTS 378
Db 382 SKAKEMLTNSVKELTS 398
```

RESULT 8

```
US-09-596-746a-60
; Sequence 60, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Datwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 399
; TYPE: PRT
; ORGANISM: ospC Chlamera
US-09-596-746a-60
```

Query Match 97.0%; Score 1802.5; DB 19; Length 399;
Best Local Similarity 97.9%; Pred. No. 4.8e-119;
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKKTTSDNAVLLAVEVEALLSSIDEIAAKAI 61
Db 26 SCNNSGKDGNTSANSADSVKGNLTETINKKTTSDNAVLLAVEVEALLSSIDEIAAKAI 85
QY 62 GKRIHONNGLDTEYNNHNSLAGAVAI STLTKOKLDGKNEGKEKIDAAKCSSEFTFNK 121
Db 86 GKRIHONNGLDTEYNNHNSLAGAVAI STLTKOKLDGKNEGKEKIDAAKCSSEFTFNK 145
QY 122 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 181
Db 146 LKEKHTDGLKGEVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAKEMLANSYKE 205
QY 182 LTPVVAESPCKPSKVNNSGKDGNTSANSADSVKGNLTETINKKTTESNAVYLAKEVE 241
Db 206 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGNLTETINKKTTESNAVYLAKEVE 262
QY 242 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAVILSNLTTRKISAIKDSGELKAE 301
Db 263 TLTSIDELAKAIGKRIKNDVSLDNEADHNSLISGAVILSNLTTRKISAIKDSGELKAE 322
QY 302 IEKAKKSEEFATAKLGHEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 361
Db 323 IEKAKKSEEFATAKLGHEHTDGLKGEVTDNNAKKAITKNNDKTGAELEKLFESYKNL 382
QY 362 SKAKEMLTNSVKELTS 378
Db 383 SKAKEMLTNSVKELTS 399
```

```
RESULT 9
US-09-596-746-30
; Sequence 30, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match
Best Local Similarity 86.8%; Score 1614.5; DB 19; Length 377;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDEIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDEIAKA 60
OY 61 IGRKHQNNGLDTEENNHNSSLGAVAIISLTIQKLDGLNEGLKEKIDAKKCEFTN 120
DB 61 IGRKHQNNGLDTEENNHNSSLGAVAIISLTIQKLDGLNEGLKEKIDAKKCEFTN 120
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
OY 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEV 240
DB 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEV 240
OY 241 ETLTSTIDELA-KAIGKKIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
DB 241 ETLTSTIDELA-KAIGKKIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
OY 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNNAKKAILTNDKTKGADELKLFESVK 359
DB 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNNAKKAILTNDKTKGADELKLFESVK 359
OY 301 EKTEENAKKCEDEFTKKLEGHAGLGIENYTDENAKKAILTTDAKDKGALEKLFRAVE 360
DB 301 EKTEENAKKCEDEFTKKLEGHAGLGIENYTDENAKKAILTTDAKDKGALEKLFRAVE 360
OY 360 NLSKAEMLTNSVKEL 376
DB 361 NLSKAEMLTNSVKEL 376
DB 361 NLSKAEMLTNSVKEL 376

RESULT 10
US-09-596-746a-30
; Sequence 30, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-30

Query Match
Best Local Similarity 87.8%; Pred. No. 9e-106;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;
```

```
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-30

Query Match
Best Local Similarity 86.8%; Score 1614.5; DB 19; Length 377;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDEIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTDSNAVLAVKEVEALLSIDEIAKA 60
OY 61 IGRKHQNNGLDTEENNHNSSLGAVAIISLTIQKLDGLNEGLKEKIDAKKCEFTN 120
DB 61 IGRKHQNNGLDTEENNHNSSLGAVAIISLTIQKLDGLNEGLKEKIDAKKCEFTN 120
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
OY 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDADAEALILKANGTKTGAEELGKLFESVEVLSKAEMLANSVK 180
OY 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEV 240
DB 181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTETISKRITTESNAVLAVKEV 240
OY 241 ETLTSTIDELA-KAIGKKIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
DB 241 ETLTSTIDELA-KAIGKKIKNDVSLDNEADHNSLSGATLISLTKKISATKDSGELK 299
OY 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNNAKKAILTNDKTKGADELKLFESVK 359
DB 300 AEIEKAKKCESEFTAKLKGHTDLGKEGVTDNNAKKAILTNDKTKGADELKLFESVK 359
OY 301 EKTEENAKKCEDEFTKKLEGHAGLGIENYTDENAKKAILTTDAKDKGALEKLFRAVE 360
DB 301 EKTEENAKKCEDEFTKKLEGHAGLGIENYTDENAKKAILTTDAKDKGALEKLFRAVE 360
OY 360 NLSKAEMLTNSVKEL 376
DB 361 NLSKAEMLTNSVKEL 376
DB 361 NLSKAEMLTNSVKEL 376

RESULT 11
US-09-596-746-82
; Sequence 82, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 400
; TYPE: PRF
; ORGANISM: OspC Chimera
US-09-596-746-82
```

Query Match 85.5%; Score 1589.5; DB 19; Length 400;
Best Local Similarity 86.7%; Pred No. 5.7e-104;
Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEIKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 25 SCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAAKAI 84
QY 62 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCESETPTNK 121
DB 85 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCESETPTNK 144
QY 122 LKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVKE 181
DB 145 LKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVKE 204
QY 182 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 241
DB 205 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 264
QY 242 TLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLITKKISAKDSGLKA 300
DB 265 TLASIDELATKAIGKIIQONGGLAVEAGHNGTLAGATTLISKLITOKLDGLKNSKLEKE 324
QY 301 EIEKAKKCESEPTAKLKGHTDLCGEVTDADAKKEAILKANGTKTGAELEKLFESVKN 360
DB 325 KIENAKKCESEPTKRLKEGHAQIGIENVYDENAKKAILITDAKDKGAAELEKLFRAVEN 384
QY 361 LSKAAKEMLTNSVKEL 376
DB 385 LAKAAKEMLANSVKEL 400
RESULT 12
US-09-596-746a-82
; Sequence 82, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 401
; TYPE: PRF
; ORGANISM: ospc Chimera
US-09-596-746a-82

Query Match 85.5%; Score 1589.5; DB 19; Length 401;
Best Local Similarity 86.7%; Pred No. 5.8e-104;
Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEIKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAAKAI 85
QY 62 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCESETPTNK 121
DB 86 GKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCESETPTNK 145
QY 122 LKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVKE 181
DB 145 LKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVKE 204

DB 146 LKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVKE 205
QY 182 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 241
DB 206 LTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 265
QY 242 TLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLITKKISAKDSGLKA 300
DB 266 TLASIDELATKAIGKIIQONGGLAVEAGHNGTLAGATTLISKLITOKLDGLKNSKLEKE 325
QY 301 EIEKAKKCESEPTAKLKGHTDLCGEVTDADAKKEAILKANGTKTGAELEKLFESVKN 360
DB 326 KIENAKKCESEPTKRLKEGHAQIGIENVYDENAKKAILITDAKDKGAAELEKLFRAVEN 385
QY 361 LSKAAKEMLTNSVKEL 376
DB 386 LAKAAKEMLANSVKEL 401

RESULT 13
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuitzen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ospc Chimera
US-09-596-746a-26

Query Match 84.2%; Score 1566; DB 19; Length 374;
Best Local Similarity 85.9%; Pred No. 2.4e-102;
Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEIKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 60
DB 1 MACNNSGKDGNTSANSADSVKGPNTLEISKTTDSNAVLAVKEVEALLSSIDEIAAKAI 60
QY 61 IGKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCESETPTN 120
DB 61 IGKRIHONNGLDTEYHNHNSLLAGAYASTLIKOKLDGLKNGELKEKIDAAKCESETPTN 120
QY 121 KKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVK 180
DB 121 KKEKHTDLCGEVTDADAKKEAILKANGTKTGAEELKLFESVEVLSKAAKEMLANSVK 180
QY 181 ELTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 240
DB 181 ELTPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEVE 240
QY 241 ETLTSIDELA-KAIGKKIKNDVSLDNEADHNSLISGAYLISNLITKKISAKDSGLKA 299
DB 238 ETLASIDELATKAIGKIIQONGGLAVEAGHNGTLAGATTLISKLITOKLDGLKNSKLEKE 297
QY 300 AIEKAKKCESEPTAKLKGHTDLCGEVTDADAKKEAILKANGTKTGAELEKLFESVKN 359
DB 298 EKIENAKKCESEPTKRLKEGHAQIGIENVYDENAKKAILITDAKDKGAAELEKLFRAVE 357

QY 360 NUSKAKEMLTNSVKEL 376
DB 358 LAKAKEMLTNSVKEL 374

RESULT 14

US-09-596-746-26
; Sequence 26, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmoist, Gerald
; APPLICANT: Dykhuisen, Dantale
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 373
; TYPE: PRN
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ospa Chimera
US-09-596-746-26

Query Match 84.0%; Score 1561; DB 19; Length 373;

Best Local Similarity 85.9%; Pred. No. 5.4e-102;

Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 2 ACNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVEALLSIDEIAAKAI 61
DB 1 ACNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVEALLSIDEIAAKAI 60
QY 62 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKLEKIDAKKCEFTTNK 121
DB 61 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKLEKIDAKKCEFTTNK 120
QY 122 LKEKHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEVLSKAKEMLTNSVKEL 181
DB 121 LKEKHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEVLSKAKEMLTNSVKEL 180
QY 182 LTPVVAESPKRPMVNNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVE 241
DB 181 LTPVVAESPKRPMVNNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVE 237
QY 242 TLLTSTIDEA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 300
DB 238 TLLTSTIDEA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 297
QY 301 EIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEV 360
DB 298 KIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEV 357
QY 361 LSKAKEMLTNSVKEL 376
DB 358 LAKAKEMLTNSVKEL 373

RESULT 15

US-09-596-746-62
; Sequence 62, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmoist, Gerald
; APPLICANT: Dykhuisen, Dantale

; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 397
; TYPE: PRN
; ORGANISM: ospc Chimera
US-09-596-746-62

Query Match 83.8%; Score 1558; DB 19; Length 397;

Best Local Similarity 85.6%; Pred. No. 9.6e-102;

Matches 322; Conservative 23; Mismatches 27; Indels 4; Gaps 2;

QY 2 ACNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVEALLSIDEIAAKAI 61
DB 25 SCNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVEALLSIDEIAAKAI 84
QY 62 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKLEKIDAKKCEFTTNK 121
DB 85 GKRIHONGDTEENNHNHNSLAGAVAI STLKOKLDGKNEGKLEKIDAKKCEFTTNK 144
QY 122 LKEKHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEVLSKAKEMLTNSVKEL 181
DB 145 LKEKHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEVLSKAKEMLTNSVKEL 204
QY 182 LTPVVAESPKRPMVNNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVE 241
DB 205 LTPVVAESPKRPMVNNNSKDGNTSANSADSEYKGNLTETINKKITSNAVLLAVEVE 261
QY 242 TLLTSTIDEA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 300
DB 262 TLLTSTIDEA-KAIGKKIKNDVSLDNEADHNSLISGAVLISLITRKISAIKDSGELKA 321
QY 301 EIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEV 360
DB 322 KIEKAKKCEFTTAKLGEHTDLGEGVTDADAKAAILKANGTKTGAELGKLFESVEV 381
QY 361 LSKAKEMLTNSVKEL 376
DB 382 LAKAKEMLTNSVKEL 397

Search completed: March 18, 2002, 10:08:43

Job time: 971 sec


```

; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel Ospc-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-974-992-7

```

```

Query Match
Best Local Similarity 33.9%; Score 631; DB 6; Length 212;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNSGKDGNT-SANSADSVKGPNTLEINKTTDSNAVLAVKEVEALLSSIDEIAAK 59
Db 17 ISCNNSGKGDGSASTNADSAKPNLTETSKITDSNAVLAVKEVEALLSSIDEIAAK 76
QY 60 AIGKIHQNNGLDTENNHNSLAGAVAISTLIKOKLDGLKN-EGLKEKIDAAKCGSEFTN 118
Db 77 AIGKIDNNNGGLAANNQNSLAGAVAISTLIKOKLDGLKN-EGLKEKIDAAKCGSEFTN 136
QY 119 TNKLEKHTDLGKEGYTDADAKKAEILKANGTKKGADELGFESVEVLSKAKEMLANS 178
Db 137 TNKLSGHALGKQDATDHAKAAILTKHATYDKAKERKDLFESVEGLKKAQVALTNS 196
QY 179 VKELTSPVVAESPKKP 194
Db 197 VKELTSPVVAESPKKP 212

```

```

RESULT 3
US-09-974-992-3
; Sequence 3, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel Ospc-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-974-992-3

```

```

Query Match
Best Local Similarity 32.1%; Score 596.5; DB 6; Length 207;
Matches 128; Conservative 25; Mismatches 38; Indels 3; Gaps 2;

```

```

QY 1 MACNSGKDGNTSANSADSVKGPNTLEINKTTDSNAVLAVKEVEALLSSIDEIAAK 60
Db 17 ISCNNSG--GDSASTNDESAKPNLTETSKITDSNAVLAVKEVEALLSSIDEI--SKA 73
QY 61 IGRKIHQNNGLDTENNHNSLAGAVAISTLIKOKLDGLKNGLKKEKIDAAKCGSEFTN 120
Db 74 IGRKIKNGDTLDNEANRNNSLNGAVEISKLITOKLSVNSELEKKEKIDKCGSEFTT 133
QY 121 KLEKHTDLGKEGYTDADAKKAEILKANGTKKGADELGFESVEVLSKAKEMLANSVK 180
Db 134 KLSHAEELGDSVODNAKKAALKTKHGTGKGAKELELETFKESLSLAKAQAALTSNVK 193
QY 181 ELTSPVVAESPKKP 194
Db 194 ELTNPVVAESPKKP 207

```

```

RESULT 4
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; NUMBER OF SEQ ID NOS: 2001-12-20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7646
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

```

```

Query Match
Best Local Similarity 9.2%; Score 171.5; DB 7; Length 1881;
Matches 100; Conservative 78; Mismatches 171; Indels 127; Gaps 17;

QY 1 MACNSGKDGNTSANSADSVKGPNTLEINKTTDSNAVLAVKEVEALLSSIDEIAAK 60
Db 1411 LAAOLETKTSNLDSTPMLEKTELEIKVKNELTEATSELTQKQNNQSLTEIEKTKAA 1470
QY 61 IGRKIHQNNGLDTENNHNSLAGAVAISTLIKOKLDGLKN-----GLKEKIDAAKCC 114
Db 1471 LTK---SSKDEYCGNQKSELQDSLKS---VASELKNFNENKYNQETSLKDEIBEKQKE 1523
QY 115 SEFTTKLEKHTDLGKEGYTDADAKKAEILK-----NG 148
Db 1524 IYVLTQLELMDRSEVKEKAMLSENSEIYIKESDKIKSLESKINSIKENSKEITTHNE 1583
QY 149 TKKGADELGFESVEVLSKAKEMLANSVKELTSPVVAESPKKPSVWNNNGKDGNTSA 208
Db 1584 QKTSLKQDIKLSODHE--SAQTQLEDKEHQKEL-----KASL-----EKHNTESA 1628
QY 209 NSADESVKGPNTLEISKKT-----TESNAVVLAVKEVEALLSSIDEIAAK---- 256
Db 1629 TSTEE--KNNQKELSEITKSLKTELKTSGDALKOSQREYTKLKNQSDTESKQOLEE 1686
QY 257 --KINDVSLDNEADHNSLISGAYLISNLTKKISAIKDSGELKA-----ETEKAK-- 306
Db 1687 LEKVKSQDTQADE-KLKGITEREIALKSELETVKKNSGLSTSELAALTKYKSLKEKEE 1745
QY 307 -----KCEEFTALKGEHNDLKG--BGVD-----DNKKAALKTNNDKTK 346
Db 1746 LQPLSGNKSKELEDYTK-----HSDISEKALKALTDLEKTKQPDSDSKKLTLENDLTS 1801

```

OY 347 GADELE-----KLFESYKNSAKAKEMLTNSYKELTS 378
DB 1802 TKKELETEKTOTSKFRNLEERKDEIYVKLNKELELLKNDNSGAKKELSEKVSLES 1857

RESULT 5

US-10-072-851-15590
; Sequence 15590, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Cair, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jlang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072.851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: fastseq for windows version 4.0
; SEQ ID NO 15590
; LENGTH: 1881
; TYPE: PRF
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-072-851-15590

Query Match 9.2%; Score 171.5; DB 7; Length 1881;
Best Local Similarity 21.0%; Pred. No. 0.0025;
Matches 100; Conservative 78; Mismatches 171; Indels 127; Gaps 17;

OY 1 MACNNSGKDGNTSANSADSVKGNLTETNKKITDSNANVLAVKEVEALLSDEIATAKA 60
DB 1411 LAAOLETKNSINDSTTMELEKTELEKVRNELTEATSELTGLKODNOSLTETIEIKTKAA 1470
OY 61 ICKKIHONNGLDPTENNHNHSLAGAVASTLTKOKLDGKNE-----GLKEKIDAKKC 114
DB 1471 LTK-----SSADLEVCGNOKSELDDLSK---VKSELKNENKYNQETTSLLKDEIEKKE 1523
OY 115 SETFNKLKEKHTDLGKECVTDADAKKALKA-----NG 148
DB 1524 IYTLQDTLKDRISEVEKEKAMLSSENSEYIVKEYSODKIKLSKINSIKENHSKEITTHNB 1583
OY 149 TTKGABEELGKLPESYEVLSKAKEMLANSVKELETSPPVAESPCKPSMYNNSGKDGNTSA 208
DB 1584 QKTSLKODIAKLSODHE--SAQTOLEDEKEMQLKEL-----KASL-----EKHNTESA 1628
OY 209 NNADESVKGNLTETSKIT-----TESNAVVLAVKREVEPTLLSIDELAKAIGK----- 256
DB 1629 TTEIE--KKNQKELSETSKLTELKTSGLDKOSOKYKTLKTKYNSPTESKLEQOLE 1686
OY 257 --KIKNDVSLNDADHNGSLISGAVLISMLTKKISAIKDSGELKA-----EIEKAK-- 306
DB 1687 LKVKVSDLOTADE-KLKGITREIALKSELETIVKNSGLSTSELALTLTKVSKLEKEKEE 1745
OY 307 -----KCBEPFAKLGKHTDLGK--EGVTD-----DNAKAILKTNNDKTK 346

DB 1746 LQFLSGNKKKELEDYIYK-----HSDISEKLMKALTFDELKEXTKQFDOSKKRKLFLDNDLTS 1801
OY 347 GADELE-----KLFESYKNSAKAKEMLTNSYKELTS 378
DB 1802 TKKELETEKTOTSKFRNLEERKDEIYVKLNKELELLKNDNSGAKKELSEKVSLES 1857

RESULT 6

US-09-708-427-15044
; Sequence 15044, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708.427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15044
; LENGTH: 1313
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1313
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1313
; OTHER INFORMATION: Ceres Seq. ID 1828627
US-09-708-427-15044

Query Match 9.2%; Score 171; DB 6; Length 1313;
Best Local Similarity 21.8%; Pred. No. 0.0016;
Matches 115; Conservative 61; Mismatches 148; Indels 204; Gaps 21;

OY 8 KDGNTSANSADSVKGNLTETNKKITDSNAV-----LLAVKEVEALLSDEIATAKA 60
DB 6 KTGIMETPRSKPSPPPRLSKLSASKSDNSASPKVHSRLVKGTETLOLQTOEDLKKA 65
OY 61 ICKKIHONNGLDPTENNHNHSLAGAVASTLTKOKLDGKNE-----LDGLK-----NEGKKEI 108
DB 66 -----DEQ-----TELKDKAKKAIIDLKESKLEVEANEKLEAL 101
OY 109 DAAKCSSEF-----TNK-----LKEKHT-DLG----- 130
DB 102 AOKKRAESEEVEKFRVLEQAGLEAVOKKDYTSKNELESIRSQHALDISALLSTTEEL 161
OY 131 ---KEGVTDADAKKAILKANGTKTGAE-----ELGKL-----FES 163
DB 162 QRYKHLSMTADAKKNAKLSHAEGATKIAELHAKEILLASELGRKALAGSKEKEAIEG 221
OY 164 VEYLSKAKE-----MLANSVKE-----LTSPPVAESP 191
DB 222 NEVYSKLEIEILLRGELEKVSILSESLKEQEGVLQVDLEPAKMAESCTNSVVEEMK 281
OY 192 KPSMYNNSGKDGNTSANSADSVKGNLTETNKKITDSNAV-----LA 236
DB 282 NKVHELEKEVEESNRSKSSASEME-----SVKKQALNHNVLHETKSDNAAKKEKIEL 336
OY 237 VKEVEFTLTSIDE-----LAKAIGKIKNDV-----SLDNEADHNGSL 274
DB 337 EKTIEAQRFDELEYGROVCAKEASKLELVLSIKSELETSOEKTRALDNNKATSN-- 395
OY 275 ISGAVLISMLTKKISAIKDSGELKAEIEKAKCSEFTAKKLGKHTDLGKEVTDNMAK 334
DB 396 -----IOMLDQRTLSLELRCRKYVEEKSCKDMESLTAL-----QEAATESSEA 441
OY 335 KATLKTNNTKTKGADLEKLFESVKNLSKAKE-----MLTNSVKELETS 378
DB 442 KATLVQCEELKNC-----ESQVDSIKLASKETNKYEKMLEDANNEIDS 486

RESULT 7
US-09-708-427-15045
; Sequence 15045, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15045
; LENGTH: 1304
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1304
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1304
; OTHER INFORMATION: Ceres Seq. ID 1828628
US-09-708-427-15045

Query Match
Best Local Similarity 9.18; Score 170; DB 6; Length 1304;
Matches 112; Conservative 60; Mismatches 136; Indels 204; Gaps 21;
QY 24 PNLTEINKKIDTSNAV-----LLAVKEVALISIDETAAKAIKKIHONNGLDTENN 76
DB 13 PRLKLSKSDSDSASPVAHSRLVKETELQTLQIQEDLKA-----DEQ-- 59
QY 77 HNGSLGAVATLILKOK---LDGLK-----NEGLEKIDAKKCSFT 118
DB 60 -----IELKKDKAKAIDDLKESKLEVEANKEKELAAQKAESFEYKFR 108
QY 119 -----TNK-----LKEKHT-DLG-----KEGYTDADAKA 142
DB 109 AVELEQGLAVOKKDVTSKNELESIRSOHALDISALSTTELORVKHELSMTADAKNK 168
QY 143 ILKANGTKTKGAE-----ELGKL-----PESEVLSKAKE----- 173
DB 169 ALSHAEATRIAEIHAKEAELILASELGRKLALGSKKEKEALEGNEIYLSKLSKSETELLRG 228
QY 174 -----MLANSYKE-----LTSVVAESPKRPSMVNNSGKGNTS 207
DB 229 ELEKVSILSSSLKEQGLVEQLKVDLEAKMAESCNTSSVEEMKNKVHLEKVEESNRS 288
QY 208 ANSADESVAKPNLITELSKITTESNAV-----LAVKEVETLLTSTDE--- 249
DB 289 KSSASESME-----SYMQLAELNHVLEHTSDNAQAKETELLEKTLEAORTDLEEYGR 343
QY 250 -----LKAIGKTKIKNV-----SLDNEADHNGSLISGAYLISNLITKIS 290
DB 344 QVCIAKEKSKLENLVESIKSELEISOEKTALDNEKAATSN-----IQNLDDPTE 396
QY 291 AIKDSGELKAELIEKAKKSEETAKLGEHTDLKEGYTDNAKALLKTNNDKTKGAE 350
DB 397 LSTELERCKVEEKSCKMESTLAL-----QEASTESSEKATLVCOEELKNC-- 446
QY 351 LEKLFESYKNISSKAKE-----MLTNSVKELTNS 378
DB 447 -ESOVDSIKLASKETNEKYEMLEDAENEIDS 477

RESULT 8
US-09-708-427-19883
; Sequence 19883, Application US/09708427
; GENERAL INFORMATION:

APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1014
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1014
; OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

Query Match
Best Local Similarity 9.08; Score 167.5; DB 6; Length 1014;
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;
QY 8 KDGNTSANSAD-----SVKGNL-----TEINKKIDTS-NAVLAAYE 45
DB 401 KHGETADSKYGLQVAELQSTLEAFYKSSSLAALNITENEKELTENLVNVSSEKKR 460
QY 46 VEALLSIDETIAK-----AIGKKIHONNGLDTENNHNSLGAAYAISTL--IKQ 94
DB 461 LEA---TVDEYSVKISSENLISIRNELNVTOG-KLESIENDLKAAGQESVMEKLS 516
QY 95 KLDGKNEGLEKEKIDAA-----KKCSEFTNK-----LKEKH 126
DB 517 AEESELEQKG--REIDETTRMELALHOSLSDSEHRLQKAMEFTSDSEASSLTERL 574
QY 127 TDL-GKEGYTDADAKEAELKANGTKTKGAEELGKLESEVYLSKAKEMLANSVKEL--- 182
DB 575 RDLBGRKTSYEOLABASGSSSLKEKLEQTLGRLLAAESVNEKIKQEPDQAKESLOSS 634
QY 183 -TSPVVAESPKR-----PSMVNNSGKGNTSANSDESVAESKGNLITELSKITTESNAV 233
DB 635 SESELLAETNNQKIKIQLEGLIGSGVEKETALRLLEALE-----RFNOKETESSDL 689
QY 234 VLAKEVEVLLISIDELA-KAIG-----KKIKN----- 260
DB 690 VEKLTQHENOIEYKRLAEASGVADTRKVELELDALSKIKNLESTIEELGAKCGGLEKES 749
QY 261 -----DVSLDNEADHNGSLISGAYLISNLITKISAIKDSGELKA-EIEKAKKCSSEFT 313
DB 750 GDIAEVNKLINLELNHGS-----EANELOTKISALEAKEQYAELEASKTTIIDLT 802
QY 314 AKL--KGEHTDLKGEYTDNAKKAELKTNNDKTKGAELEKLEESYK-NLSKAKEMLT 370
DB 803 KQLTSEGEKLOSQLEKRLVAASKEVSLESH-----FEELKLTSEVKAOLKEVENAAT 856
QY 371 NSVK--ELTS 378
DB 857 ASVKAVALTS 866

RESULT 9
US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09

```

RESULT 10
US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

```

```

RESULT 11
; Sequence 5883, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zykien, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

```

;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5883
;; LENGTH: 837
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-5883

Query Match 8.7%; Score 162.5; DB 6; Length 837;
Best Local Similarity 18.7%; Pred. No. 0.0033;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSEYKGP--NLTETINK-----KITDSNAVLAVKEVEALLSS 52
DB 350 NGNQNVANAKTTAKNALNLITSIINNAQKEALKSQIEGATTVAGVNOVSTTASELNTAMSN 409
QY 53 I-----DEIAAKAIGKKIHQNNGLDTENNHNSILAGAVAISTLIKOKLDGLKNGEKER 107
DB 410 LONGINDEAATKA--ALNGTONLEKAKOHANTAIIDL-----SHLTNAQKEALKQOL 458
QY 108 IDAAKCEFTFTNKLKEKHTDGLKEGVTDAADAKAELKANGTKTKGAELKLFESVEVL 167
DB 459 VOOSTTVAAGQNEQKANNVDAAMDKLRSIDNFTTKONONTDASQKKKAYNNAYTT 518
QY 168 SRAAKEMLANSVKELTSPVVAESPCKPSVNNNGKDGNTSANSADSEYKGPMLTETISKI 227
DB 519 AGCI-----IDQTSPTL-----DPTVINOAGQVSTTKNALNGN--ENLEAAKQQA 563
QY 228 TESNAVLAKEVEFTLLTSIDELAKAIGKKIKNDVSLDNEADHNSILSGAVLISMLITK 287
DB 564 SOS-----LGSIDLNNNAQKQVTDQINGAHTVDEANQIKONQNLNTAMGN 610
QY 288 KISAIDSEELKAEI-----EKAKK-----CSEFTAKLKGEHTDGL--KEGVTDN 332
DB 611 LKQALADKATKATVNTFDADQAKQOAVNTAVTNMENIISKANGNATQAEVEQAIKQVN 670
DB 333 AKKALITNNDKTKGADELEKLFESYKMLSKAAKEMLINSVEELTS 378
DB 671 AAKQALNGNANVQAHKDEATALINSNDLNQAKQKQVONAYTT 716

RESULT 12
US-10-072-851-5883
; Sequence 5883, Application US/10072851
; GENERAL INFORMATION:

;; APPLICANT: Carr, Grant J.
;; APPLICANT: Xu, H. Howard
;; APPLICANT: Foulkes, J. Gordon
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Kari L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Roemer, Terry
;; APPLICANT: Jiang, Bo
;; APPLICANT: Boone, Charles
;; APPLICANT: Bussey, Howard
;; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
;; FILE REFERENCE: ELITRA.028A
;; CURRENT APPLICATION NUMBER: US/10/072,851

;; CURRENT FILING DATE: 2002-02-08
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; NUMBER OF SEQ ID NOS: 15811
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5883
;; LENGTH: 837
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-10-072-851-5883

Query Match 8.7%; Score 162.5; DB 7; Length 837;
Best Local Similarity 18.7%; Pred. No. 0.0033;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSEYKGP--NLTETINK-----KITDSNAVLAVKEVEALLSS 52
DB 350 NGNQNVANAKTTAKNALNLITSIINNAQKEALKSQIEGATTVAGVNOVSTTASELNTAMSN 409
QY 53 I-----DEIAAKAIGKKIHQNNGLDTENNHNSILAGAVAISTLIKOKLDGLKNGEKER 107
DB 410 LONGINDEAATKA--ALNGTONLEKAKOHANTAIIDL-----SHLTNAQKEALKQOL 458
QY 108 IDAAKCEFTFTNKLKEKHTDGLKEGVTDAADAKAELKANGTKTKGAELKLFESVEVL 167
DB 459 VOOSTTVAAGQNEQKANNVDAAMDKLRSIDNFTTKONONTDASQKKKAYNNAYTT 518
QY 168 SRAAKEMLANSVKELTSPVVAESPCKPSVNNNGKDGNTSANSADSEYKGPMLTETISKI 227
DB 519 AGCI-----IDQTSPTL-----DPTVINOAGQVSTTKNALNGN--ENLEAAKQQA 563
QY 228 TESNAVLAKEVEFTLLTSIDELAKAIGKKIKNDVSLDNEADHNSILSGAVLISMLITK 287
DB 564 SOS-----LGSIDLNNNAQKQVTDQINGAHTVDEANQIKONQNLNTAMGN 610
QY 288 KISAIDSEELKAEI-----EKAKK-----CSEFTAKLKGEHTDGL--KEGVTDN 332
DB 611 LKQALADKATKATVNTFDADQAKQOAVNTAVTNMENIISKANGNATQAEVEQAIKQVN 670
QY 333 AKKALITNNDKTKGADELEKLFESYKMLSKAAKEMLINSVEELTS 378
DB 671 AAKQALNGNANVQAHKDEATALINSNDLNQAKQKQVONAYTT 716

RESULT 13
US-09-815-242-13080
; Sequence 13080, Application US/09815242
; GENERAL INFORMATION:

;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Kari L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931

;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13080
;; LENGTH: 875
;; TYPE: PRF
;; ORGANISM: Staphylococcus aureus
US-09-815-242-13080

Query Match 8.7%; Score 162.5; DB 6; Length 875;
Best Local Similarity 18.7%; Pred. No. 0.0035;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSVKGP--NLTEINK-----KITDSNAVLAVKEVEALLSS 52
DB 387 NGNONVANAKTTAKNALNLTSTINNOKKALQSIEGATTVAGVQVSTTASELNTAMSN 446
QY 53 I-----DEIAKAIGKKIHONNGIDTENNHGSLLAGAVASTLIKOKLDGKNEGKLEK 107
DB 447 LONGINDEAATKA--ALNGTONLEKAKQHANTAIDGL-----SHLTNAOKKALKOL 495
QY 108 IDAKKSEFTTNKLEKHNDLGEKGYTDADAKFAILLKANGTKTKGAELGKLEESYEV 167
DB 496 VQOSTTVAEAGNEOKANNVDAADKROSTADNATTKONQNTDASONKKDAYNNAVTT 555
QY 168 SKAKKEMLANSVKELTSPVVAESPSPKPSMVNSGDKGNTSANSADSVKGPNTLEISKI 227
DB 556 AAGI-----IDGTTSPTL-----DPTVINQAAQVSTTKNALNGN---ENLEAAQQA 600
QY 228 TESNAVAVLAVEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTK 287
DB 601 SOS-----LGSIDLNNNAOKQVTTDQINGAHVYDEANOIQKONQONNTAMGN 647
QY 288 KISAIKDSGELKAEI-----EKAKK-----CSEFTAKLKEGHDGLG--KEGYDDN 332
DB 648 LKQAIADKATKATVNTDADQAKQOAVNTAVTNAENITISANGNATQAEVEAOKOVN 707
QY 333 AKKAILTKNDKTGADELKLFESVKNLSKRAKEMLTNSVKELTS 378
DB 708 AAKQALNGNANVQAKDEATLINSNDLNOQADALKQOVQNMNTT 753

RESULT 14
US-10-072-851-13080
; Sequence 13080, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13080

;; LENGTH: 875
;; TYPE: PRF
;; ORGANISM: Staphylococcus aureus
US-10-072-851-13080

Query Match 8.7%; Score 162.5; DB 7; Length 875;
Best Local Similarity 18.7%; Pred. No. 0.0035;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

QY 9 DGNTSANSADSVKGP--NLTEINK-----KITDSNAVLAVKEVEALLSS 52
DB 387 NGNONVANAKTTAKNALNLTSTINNOKKALQSIEGATTVAGVQVSTTASELNTAMSN 446
QY 53 I-----DEIAKAIGKKIHONNGIDTENNHGSLLAGAVASTLIKOKLDGKNEGKLEK 107
DB 447 LONGINDEAATKA--ALNGTONLEKAKQHANTAIDGL-----SHLTNAOKKALKOL 495
QY 108 IDAKKSEFTTNKLEKHNDLGEKGYTDADAKFAILLKANGTKTKGAELGKLEESYEV 167
DB 496 VQOSTTVAEAGNEOKANNVDAADKROSTADNATTKONQNTDASONKKDAYNNAVTT 555
QY 168 SKAKKEMLANSVKELTSPVVAESPSPKPSMVNSGDKGNTSANSADSVKGPNTLEISKI 227
DB 556 AAGI-----IDGTTSPTL-----DPTVINQAAQVSTTKNALNGN---ENLEAAQQA 600
QY 228 TESNAVAVLAVEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTK 287
DB 601 SOS-----LGSIDLNNNAOKQVTTDQINGAHVYDEANOIQKONQONNTAMGN 647
QY 288 KISAIKDSGELKAEI-----EKAKK-----CSEFTAKLKEGHDGLG--KEGYDDN 332
DB 648 LKQAIADKATKATVNTDADQAKQOAVNTAVTNAENITISANGNATQAEVEAOKOVN 707
QY 333 AKKAILTKNDKTGADELKLFESVKNLSKRAKEMLTNSVKELTS 378
DB 708 AAKQALNGNANVQAKDEATLINSNDLNOQADALKQOVQNMNTT 753

RESULT 15
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5835

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:42 ; Search time 39.62 Seconds

(without alignments)
349.806 Million cell updates/sec

Title: US-09-596-746A-28

Sequence score: 1859
1 MACNNSGKDGNTSANSADES.....KNLSKAKEMLTNSVKELTS 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Partial number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008
Listing first 45 summaries

Database : SWISSPROT_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	50.8	210	1	OSCL_BORBU
2	631	33.9	212	1	OSCL_BORBU
3	397.5	21.4	214	1	VM24_BORHE
4	397	21.4	215	1	VM03_BORHE
5	179	9.6	1509	1	MYSN_ACACA
6	173.5	9.3	1957	1	YD86_SCHPO
7	168.5	9.1	483	1	M6_STRPY
8	168.5	9.1	564	1	M12_STRPY
9	167.5	9.0	998	1	SCA4_RICAF
10	166.5	9.0	1251	1	RBP2_PLAVB
11	165.5	8.9	1938	1	MYSL_PUVR
12	164	8.8	1790	1	USO1_YEAST
13	163.5	8.8	2116	1	MYSL_YEAST
14	163	8.8	3210	1	CENF_HUMAN
15	157.5	8.5	1966	1	MYSL_CAEEL
16	157	8.4	1786	1	LMBL_MOUSE
17	155.5	8.4	492	1	M5_STRPY
18	154	8.3	1427	1	SCA4_RICAF
19	153.5	8.3	1427	1	REST_HUMAN
20	153.5	8.3	2869	1	RBP1_PLAVB
21	149.5	8.0	1935	1	MYSL_CYPFA
22	149	8.0	1011	1	SCA4_RICAF
23	148.5	8.0	1189	1	SC11_CHICK
24	148	8.0	866	1	MYSP_SCHJA
25	148	8.0	875	1	Z1P1_YEAST
26	148	8.0	2017	1	MYSL_DROME
27	148	8.0	3110	1	LM2_HUMAN
28	147.5	7.9	705	1	YNP9_CAEEL
29	147.5	7.9	1639	1	LMG1_DROME
30	147	7.9	991	1	SCA4_RICAF
31	147	7.9	1500	1	SSP5_STRGN
32	147	7.9	1961	1	MYH9_RAT
33	146.5	7.9	967	1	KINH_LOLPE

34	146	7.9	679	1	YKR9_YEAST	P34237 saccharomyc
35	145.5	7.8	1786	1	LMB1_HUMAN	P07942 homo sapien
36	145	7.8	866	1	MYSP_SCHJA	P06198 schistosoma
37	145	7.8	1010	1	SCA4_RICPA	O94175 rickettsia
38	144.5	7.8	1164	1	BAG_STRAG	P27951 streptococc
39	144.5	7.8	1433	1	REST_CHICK	O42184 gallus gall
40	144.5	7.8	1453	1	Y373_BOVIN	O91423 bos taurus
41	144.5	7.8	2349	1	TPR_HUMAN	P12270 homo sapien
42	144	7.7	1713	1	LM2_HUMAN	O16787 homo sapien
43	143.5	7.7	1935	1	MYH7_HUMAN	P12883 homo sapien
44	143	7.7	775	1	YHGE_BACSU	P32399 bacillus su
45	143	7.7	3672	1	LM2_CAEEL	O21313 caenorhabdi

ALIGNMENTS

RESULT	1	STANDARD	PRT	210 AA.
OSCL_BORBU	007337			
AC	007337			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB19.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID=139;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=93268136; PubMed=8098841;			
RA	Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,			
RA	Will G., Wilske B.;			
RT	"Genetic heterogeneity of the genes coding for the outer surface			
RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.";			
RL	Med. Microbiol. Immunol. 182:37-50(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=9329332; PubMed=8478108;			
RA	Wilske B., Preac-Mursic V., Jauris S., Soutschek E.,			
RA	Schwab E., Wanner G.;			
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant			
RT	major outer surface protein of Borrelia burgdorferi.";			
RL	Infect. Immun. 61:2182-2191(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=94041630; PubMed=8225587;			
RA	Pedula S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;			
RT	"Molecular characterization and expression of p23 (OspC) from a North			
RT	American strain of Borrelia burgdorferi.";			
RL	Infect. Immun. 61:5097-5105(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=96025162; PubMed=7494039;			
RA	Eukunaga M., Hamase A.;			
RT	"Outer surface protein C gene sequence analysis of Borrelia			
RT	burgdorferi sensu lato isolates from Japan.";			
RL	J. Clin. Microbiol. 33:2415-2420(1995).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE=98065943; PubMed=9403685;			
RA	Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,			
RA	Iathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,			
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,			
RA	Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,			
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,			

```

RA  Utterback T., Matthey L., McDonald L., Artlach P., Bowman C.,
RA  Gairland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA  Smith H.O., Venter J.C.;
RT  "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT  burgdorferi";
RL  Nature 390:580-586(1997).
CC  -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC  -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC  ANCHOR.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X69596; CAA49306.1; -
DR  EMBL: U01894; AAA16058.1; -
DR  EMBL: D49497; BAA08457.1; -
DR  EMBL: AE000792; AAC66329.1; -
DR  TIGR: BB819; -
DR  InterPro: IPR001800; Lipoprotein_6.
DR  Pfam: PF01441; Lipoprotein_6; 1.
DR  Prodom: PD001149; Lipoprotein_6; 1.
DR  PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW  Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW  Complete proteome.
FT  SIGNAL 1 18 BY SIMILARITY.
FT  CHAIN 19 210 OUTER SURFACE PROTEIN C.
FT  LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ  SEQUENCE 210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match 50.8%; Score 945; DB 1; Length 210;
Best Local Similarity 97.9%; Pred. No. 2,6e-36;
Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETNKKTDSNAVLAVKEVEALLSSIDEIAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETNKKTDSNAVLAVKEVEALLSSIDEIAKA 76
QY 61 IGGKIHQNNGLDTENNHNSILAGAVASTLLKOKLDGKNGEIKKEKIDAAKCESTFN 120
DB 77 IGGKIHQNNGLDTENNHNSILAGAVASTLLKOKLDGKNGEIKKEKIDAAKCESTFN 136
DB 121 KIKKEKHTDLGKGVYDADAKKAAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 180
DB 137 KIKKEKHTDLGKGVYDADAKKAAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 196
QY 181 ELTSPVVAESPKRP 194
DB 197 ELTSPVVAESPKRP 210

RESULT 2
OSC2_BORBU STANDARD; PRT; 212 AA.
AC Q08137;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (PC).
GN OSpC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=PRO;
RA MEDLINE=92219995; Pubmed=1560779;
RA Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

```

```

RA  Soutschek E.;
RT  "Molecular analysis and expression of a Borrelia burgdorferi gene
RT  encoding a 22 kDa protein (PC) in Escherichia coli.";
RL  Mol. Microbiol. 6:503-509(1992).
RN [2]
RP SEQUENCE OF 1-205 FROM N.A.
RC STRAIN=DR26;
RA MEDLINE=94075528; Pubmed=8253951;
RA Theisen M., Frederiksen B., Iebach A.M., Yunst J., Hansen K.;
RT "Polymorphism in ospC gene of Borrelia burgdorferi and
RT immunoreactivity of OspC protein: implications for taxonomy and for
RT use of OspC protein as a diagnostic antigen.";
RL J. Clin. Microbiol. 31:2570-2576(1993).
CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL: X62162; CAA44093.1; -
DR  EMBL: X73624; CAA52003.1; -
DR  InterPro: IPR001800; Lipoprotein_6.
DR  Pfam: PF01441; Lipoprotein_6; 1.
DR  Prodom: PD001149; Lipoprotein_6; 1.
KW  Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT  SIGNAL 1 18 BY SIMILARITY.
FT  CHAIN 19 212 OUTER SURFACE PROTEIN C.
FT  LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ  SEQUENCE 212 AA; 22499 MW; C206C231FBE2E7D4 CRC64;

Query Match 33.9%; Score 631; DB 1; Length 212;
Best Local Similarity 69.9%; Pred. No. 2,6e-22;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETNKKTDSNAVLAVKEVEALLSSIDEIAK 59
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETNKKTDSNAVLAVKEVEALLSSIDEIAK 76
QY 60 AIGKIHQNNGLDTENNHNSILAGAVASTLLKOKLDGKN-EGKKEKIDAAKCESTFN 118
DB 77 AIGKIHQNNGLDTENNHNSILAGAVASTLLKOKLDGKN-EGKKEKIDAAKCESTFN 136
QY 119 TNKIKKEKHTDLGKGVYDADAKKAAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 178
DB 137 TNKIKSGHADLGRQDADTDHAKKAAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 196
QY 179 VKELTSPVVAESPKRP 194
DB 197 VKELTSPVVAESPKRP 212

RESULT 3
VM24_BORBU STANDARD; PRT; 214 AA.
AC P32778;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN VMP24.
OS Borrelia hermsii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-SSP. HSI SEROTYPE 24;
RX MEDLINE-93133110; PubMed-1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids
are highly polymorphic."
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04786; AAA22964.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE
FT LIPID 19 19 N-ACYL DIGLICERIDE (PROBABLE).
FT SEQUENCE 214 AA; 22541 MW; F1583F510246FC7 CRC64;

```

```

Query Match 21.4%; Score 397.5; DB 1; Length 214;
Best Local Similarity 44.8%; Pred. No. 6.7e-12;
Matches 91; Conservative 37; Mismatches 58; Indels 17; Gaps 5;

```

```

OY 1 MACNSGKCGNTSANSADSVKGP----NLTEINKKITDSNAVLAVKEVALLSIDET 56
DB 17 MCGNNGGPE-----LKSDEVAKSDGTVLDLAKISKIKTIDSDPAASVKEVHTLVKSIDEL 71
OY 57 AKAATGKTIHONNGLDTEENNNGSLAGAVAI STLKOKL-----DGLKNEGLKEKIDA 110
DB 72 -AKAATGKTIHONNGLDTEENNNGSLAGAVAI STLKOKL-----DGLKNEGLKEKIDA 110
OY 111 AKKSEFTFNKLEKHTDLEKGVTDADAKKALIKANGTKTGAEEGLKFESEVEVLSKA 170
DB 130 VKSKAEAFPLNKLKDGHTLKGKQASDDDTKKAIKKQNSDKTKGASELEALNTAVDALLKA 189
OY 171 AKEMLANSYKELTSPVVAESPKK 193
DB 190 AEGVEAAIKELTAPVAKKPSQ 212

```

```

RESULT 4
VMP3 BORNE
ID VMP3 BORNE STANDARD; PRT; 215 AA.
AC 002448;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN VMP3.
OS Borrelia hermsli.
OC Bacteriia; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_Taxid=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SSP. HSI SEROTYPE 3;
RX MEDLINE-93133110; PubMed-1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids

```

```

RT are highly polymorphic."
RL Mol. Microbiol. 6:3299-3311(1992).
CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: STRONG, TO VMP24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04789; AAA22967.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
FT LIPID 19 19 N-ACYL DIGLICERIDE (PROBABLE).
FT SEQUENCE 215 AA; 23139 MW; 684C74D35F87C771 CRC64;

```

```

Query Match 21.4%; Score 397; DB 1; Length 215;
Best Local Similarity 45.3%; Pred. No. 7.1e-12;
Matches 92; Conservative 35; Mismatches 60; Indels 16; Gaps 5;

```

```

OY 1 MACNSGKCGNTSANSADSVKGP----NLTEINKKITDSNAVLAVKEVALLSIDET 56
DB 17 MCGNNGGPE-----LKSDEVAKSDGTVLDLAKISKIKTIDSDPAASVKEVHTLVKSIDEL 71
OY 57 AKAATGKTIHONNGLDTEENNNGSLAGAVAI STLKOKL-----DGLKNEGLKEKIDA 110
DB 72 -AKAATGKTIHONNGLDTEENNNGSLAGAVAI STLKOKL-----DGLKNEGLKEKIDA 110
OY 111 AKKSEFTFNKLEKHTDLEKGVTDADAKKALIKANGTKTGAEEGLKFESEVEVLSKA 170
DB 131 VKKSEAFYTVKSKHTDLAKKEGVTDAAKSAALIVDGTDKGAALIKLNTAIDELKA 190
OY 171 AKEMLANSYKELTSPVVAESPKK 193
DB 191 ANDAVETVIKELTAPVAKKPSQ 213

```

```

RESULT 5
MYSN ACACA
ID MYSN ACACA STANDARD; PRT; 1509 AA.
AC P05659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_Taxid=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87308395; PubMed-3040773;
RA Hammer J.A. II, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail."
RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

```

CC LIGHT CHAIN SUBUNITS (MHC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (MC-2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
 CC MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Y00624; CAA68663.1; -
 DR PIR: A27224; A27224.
 DR HSSP: P08799; LMND.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IO; 1.
 DR Myosin: Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
 KW Methylation; Alkylation; Phosphorylation; Multigene family.
 FT DOMAIN 1 789
 FT MYOSIN HEAD-LIKE.
 FT 790 819
 FT COILED COIL (POTENTIAL).
 FT 848 1509
 FT ALPHA-HELICAL TAILPIECE (S2).
 FT 1227 1252
 FT HINGE.
 FT 1253 1509
 FT LIGHT MEROMYOSIN (LMN).
 FT 1253 1482
 FT ALPHA-HELICAL TAILPIECE (LMN).
 FT 1483 1509
 FT NONHELICAL TAILPIECE.
 FT 182 189
 FT ATP.
 FT 660 682
 FT ACTIN-BINDING.
 FT 766 780
 FT ACTIN-BINDING.
 FT 133 133
 FT METHYLATION (TR1-) (POTENTIAL).
 FT 700 700
 FT ALKYLATION (SH-1) (POTENTIAL).
 FT 1489 1489
 FT PHOSPHORYLATION.
 FT 1494 1494
 FT PHOSPHORYLATION.
 FT 1499 1499
 FT PHOSPHORYLATION.
 SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 9.6%; Score 179; DB 1; Length 1509;
 Best Local Similarity 24.9%; Pred. No. 0.33;
 Matches 102; Conservative 69; Mismatches 146; Indels 92; Gaps 20;

QY 28 ELNKKITSDNAVLAVKVEEA---LSSIDEIAKAIGKIHQ--ANGLDTENNHNSL 81
 DB 934 ELQERTSANDLEQKRLLEAKGELKAKSLDE--EENKRALDEANTKVESENELQDK 990
 QY 82 LAGAAVAITLIKQKLDGKNGLEKIDK---AKKSEFTFNKKEKHTDJKGKGVTDK 137
 DB 991 YEDAAHAHDSKKKEEDLSRE-LRETKDALADANISSELTLSKLT--KTERGADVANEL 1047
 QY 138 -DAEAILKANGTGTGAKEELGKLFESYEVLSKAAKEMLSKVELTSPV----- 186
 DB 1048 DDVATATKQLEKTKKSEELAEQTRAOLEE-EKSGKEAASSKAKQLOQOQLEADARSEVSL 1106
 QY 187 -----VASPKKPMVNNNG-----KDGNTSANSADSVK--PNLTETSKKIT--ES 230
 DB 1107 KSKLSAEKSLKTAQDNRLDQLEDERTVVRANVQKOKALAKLTLEDDVATALDQK 1166

QY 231 NAVLAKVEYETLLTSLDEL-----AKAIGKINDVSLDNEADHNSLISGAVL 280
 DB 1167 NA---AAQAKTLTKTQVDEETKRLLEAEASAAKLEKERN--ALDEVAQ----- 1210
 QY 281 ISNLTTRKISAKDSG-----BLKAEIEKAK--CSEFTFAKLGKHTDGLK 325
 DB 1211 ---LADDAEDSDGAQQRRLNTRISLOQSELENAKPTGASSEYKKEGELERLEE 1266
 QY 326 EGYTDNKAKKALITKNDTKGADELEKLFESYKNSLKAEMELTNSYK 374
 DB 1267 ELTLAQEAARAER-NLDKAN--LELELRQEDADDAARDNDKLVQDNK 1312

RESULT 6
 ID YD86.SCHPO STANDARD; PRT; 1957 AA.
 AC 010411:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOTHETICAL. 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME 1.
 GN SPAC1F3.06C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=972;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z70690; CAA94624.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 9.3%; Score 173.5; DB 1; Length 1957;
 Best Local Similarity 22.8%; Pred. No. 0.76;
 Matches 108; Conservative 73; Mismatches 148; Indels 145; Gaps 18;

QY 8 KDGNTSANSADK---SVKGNLTETIRKKTDSNAVLAV-KEVALLSIDET----- 56
 DB 360 KDSRTINSQLEEMVYLKLSNRT-IHSQLTDAESKLSFSEQENKSLKSIDETQYONNLSSK 418
 QY 57 -----AKAIGKIKHONNGLDTEN-----NHNSGL 82
 DB 419 DKWAKOVSSQLEEARSLAHATGLAEINSEKDPONKKIKDEKTEQDLKACLNSSNEL 478
 QY 83 AGAYALSTLIKQKLDGKNGLEKIDAKKCEFTFKL-----KEKH----- 126
 DB 479 KER---SALIDKKDDELNN--LREQIKQKQVSESTOSLSLOPDLINERKKHVEYSQ 533
 QY 127 -----TDLKEGVTADAKAEALIKANGTKTGADELGLKLFESYEVLSKAA 171
 DB 534 LNEIKGELQTEISENHEHLSQSLTAAKKAVALTN-----NELSEKNSLDTLQNAF 586
 QY 172 KEMLANSKVELTSPVVAASPKKPMVNNNGKDGNTSANSADSVKGNLTETSKKITESN 231
 DB 587 QEKLAQSVQWL-----KENQNFSSLDTSFKKLNEHQELEN-----NHQTTTQKLDTS 636
 QY 232 AVLAVLAV-----EYETLLTSLDELAKAIGKIKINDVSLDNEADHNS 273
 DB 637 SKIQQLQLERANFQKESTLSDENNDLRTKLKLEESNKSILKQEDVDLE----- 688

OY 274 LISGAVLISLITKTSIAIKDSEGLKAEIEAKKCESEFTAKLGEHTDGRK-----G 327.
 DB 689 -----KNIOTLKEDLRKSEBALKRFSKLEAKNL-REVIDNKGHEHTEAQRNDLHSS 739
 OY 328 VTDDNKKAKLITKNNDKTKGADELKLFESVKNL-----SKAAKEMLTNSVKELT 378
 DB 740 LSDAKRTNML--SSELTNSSEDKRLTANVERLTODSKRMMKSFSLVNSYOS 791

RESULT 7
 M6_STRPY ID M6_STRPY STANDARD: PRT: 483 AA.
 AC P08089;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE M PROTEIN, SEROTYPE 6 PRECURSOR.
 GN EMM6.
 OS Streptococcus pyogenes.
 NC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 NC Streptococcus.
 NCBI_TaxID=1314;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86111835; PubMed=3511046;
 RA Hollingshead S.K., Fischetti V.F., Scott J.R.;
 RT "Complete nucleotide sequence of type 6 M protein of the group A
 RT Streptococcus. Repetitive structure and membrane anchor."
 RL J. Biol. Chem. 261:1677-1686(1986).
 RN (2)
 RP SEQUENCE OF 43-122 FROM N.A.
 RX MEDLINE=85166224; PubMed=3885219;
 RA Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
 RT "Relationship of M protein genes in group A streptococci."
 RT Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
 CC PHAGOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
 CC -1- SIMILARITY: TO OTHER M PROTEINS.
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M1338; AAA26920.1; -
 DR PIR: A26297; A26297.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR003345; M_repeat.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 DR Pfam: PF02370; M; 9.
 DR PRINTS: PR00015; GP0SANCHOR.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
 KW Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
 KW Transmembrane; Coll; Signal.
 FT SIGNAL 1 42
 FT CHAIN 43 483 M PROTEIN, SEROTYPE 6.
 FT DOMAIN 43 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 477 MEMBRANE ANCHOR.
 FT DOMAIN 478 483 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 69 138 10 X 7 AA TANDEM REPEATS.
 FT DOMAIN 157 269 4.5 X 25 AA TANDEM REPEATS.
 FT DOMAIN 279 347 TWO DIRECTLY REPEATED 27 AMINO ACID
 FT DOMAIN 348 411 BLOCKS SEPARATED BY 15 AMINO ACIDS.
 FT HYDROPHILIC.

FT DOMAIN 412 448 GLY/PRO-RICH (CELL WALL-SPANNING).
 FT DOMAIN 449 454 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
 FT PROTEINS.
 SQ SEQUENCE 483 AA; 53472 MW; 68f87f28db534448 CRC64;
 Query Match 9.1%; Score 168.5; DB 1; Length 483;
 Best Local Similarity 22.9%; Pred. No. 0.27;
 Matches 91; Conservative 57; Mismatches 153; Indels 97; Gaps 15;
 OY 5 NSGKDGNTSANSADSEYKGPNTLEINKKIDTSDNVLAVEEVALSSIDEIAKAIGK 64
 DB 73 NNDKLTENNNDLQD---KNLTENKNLTDQKNLTENK-----NL 112
 OY 65 IHONNGIDTENNHNHNSLAGAVAISTLIKORLDGKNKGKIDAKKSEFTTNKLE 124
 DB 113 TDONKNLTENK---ELKAEENRLTT-----ENKGLTKKISEAE---EEANKERE 157
 OY 125 KHTDICK-EGVTADAKAEALITKANGTKTGAELGKLFESVE--VLSKAKEM----- 174
 DB 158 NKEAIGTLKTLDETETVYDKIAKQESKFT---ETIGLTKTLDETETVYDKIAKQESKFTTG 213
 OY 175 -LANSVKELTSPVAESPKKPPSMVNSGKDGNTSANSADSEYKGPNTLEISKITTESNAV 233
 DB 214 TLKKTIDETVYDKIAKQESKFT-----GTLKTLDETETVYDKIAKQESKFTTG 261
 OY 234 VLAKEVEYELLTISDELAKAIGKIKNDVSLDNEADHNSLISGAVLISLITKTSIAIK 293
 DB 262 IGALKQELAKDKGNKVSASRKGLRDLASREA-----KKQVEK 302
 OY 294 DSGELKAEIEKAKKCESEFTAKLGEHTDGRKCEVTDNKKAKLITKNNDKTKGADLEK 353
 DB 303 DLNLTAELDKVKEEKOISDASRGRLRDASREAKKQVEKA-LEBANSKLALALEKLNK 361
 OY 354 LFESVKNL-----SKAAKEMLTNSVKELT 376
 DB 362 ELEBSKLTKEKAEALQAKLEAKKALKQALANQAEEL 399

RESULT 8
 M12_STRPY ID M12_STRPY STANDARD: PRT: 564 AA.
 AC P19401;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).
 GN EMM12.
 OS Streptococcus pyogenes.
 NC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 NC Streptococcus.
 NCBI_TaxID=1314;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX SMRAIN=CS24 / Serotype M12;
 RX MEDLINE=88058777; PubMed=2445730;
 RA Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;
 RT "Streptococcus pyogenes type 12 M protein gene regulation by upstream
 RT sequences."
 RL J. Bacteriol. 169:5633-5640(1987).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
 CC PHAGOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
 CC -1- SIMILARITY: TO OTHER M PROTEINS.
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M18269; AAA88573.1; -
 DR PIR; A60115; A60115
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR003345; M_repeat.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF02370; M; 9.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
 DR Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
 KW Transmembrane; Coiled coil; Signal.
 FT SIGNAL 1 41
 FT CHAIN 1 41
 FT DOMAIN 42 564 M PROTEIN, SEROTYPE 12.
 FT TRANSMEM 42 550 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 551 554 MEMBRANE ANCHOR.
 FT DOMAIN 44 505 COILED COIL (POTENTIAL).
 FT DOMAIN 505 541 GLY/PRO-RICH (CELL WALL-SPANNING).
 FT DOMAIN 542 547 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
 FT PROTEINS.
 FT NON_TER 564 564
 FT SEQUENCE 564 AA; 62904 MW; 5F1549DACAA77B46 CRC64;

Query Match 9.1%; Score 168.5; DB 1; Length 564;
 Best Local Similarity 22.9%; Pred. No. 0.32;
 Matches 97; Conservative 59; Mismatches 166; Indels 101; Gaps 17;

QY 10 GNTSANSADSEYKGNLTETINKKTTSDNAVLAV-KEVEALLSSIDETAKAIGKTHON 68
 DB 115 GLGIDNADLNK--ITELKSEVEKNVDSQIKKELEKENDIQ-----FGREYHAA 165
 QY 69 NGIDTENNHNHNSLLAGAVAI STLKOKLDGKNGLEKIDAAKCSFTTNKLEKHTD 128
 DB 166 DL--RHKEIIEKEVVIS--KLNG-ELDPLKOKVD-----ETDRNLQEKOKV 209
 QY 129 LGKE--GYTDADAKKEAI-LKANGTKTGAEELGKLPESVEVSKAAK--EMIANVKEEL 182
 DB 210 LSLDQALAVTKENAKKDFELALGHOLADKEYNAKIAELESKIDAKKPELALAGHQA 269
 QY 183 TSPVAESPKKPMVN-----NSGKDGNTSANSADSEYKGNLTETISK-----KI 227
 DB 270 HNEQALAEKDOIKOLEBOKOILDASRGATARDLEAVROAKKATEALNMLKELKV 329
 QY 228 TESNAVLAVK-----EVEETLTSIDE--LAKAIGKIKNDVSLNDEA 268
 DB 330 TEOKOILDASRGATARDLEAVRKSKQOYEAALKOLEBOKITSEASRGGLRDLDTISREA 389
 QY 269 DHNGSLISGAVYLSNLTIKKISAIKSGELAEIEKAKKCSSEFTAKLGEHTDLGKREGV 328
 DB 390 -----KQOVERDLNLTAEELDKVEEKOISDASRGGLRDLDTISREA 430
 QY 329 TDNNAKAILKTNDKTGADELEKLFESVKL-----SKAKEMLTNSV 373
 DB 431 AKKOVEKA-LEFANSKRLALEKLNLDSESKLLEKKAELQANLEAEAKALKLEOLAKOA 489
 QY 374 KEL 376
 DB 490 EEL 492

RESULT 9
 SCA4_RICAK STANDARD; PRT; 998 AA.
 ID SCA4_RICAK
 AC 09AIX9:
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE ANTIGENIC HEAT-STABLE 120 KDA PROTEIN (Pst120) (120 KDA ANTIGEN)
 GN SCA4 OR D.
 OS Rickettsia akari.

CC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiidae; Rickettsia.
 CC MBL_TaxID=786;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sekeyova Z., Roux V., Raoult D.;
 RT "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the
 RT 'gene D' coding for an intracytoplasmic protein."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLY).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF213016; AAK30691.1; -
 KW Antigen.
 FT NON_TER 1 1
 FT NON_TER 998 998
 FT SEQUENCE 998 AA; 109328 MW; FCEB43AC62DE5BD5 CRC64;

Query Match 9.0%; Score 167.5; DB 1; Length 998;
 Best Local Similarity 20.7%; Pred. No. 0.67;
 Matches 102; Conservative 69; Mismatches 140; Indels 181; Gaps 22;

QY 5 NSG-----KDGNTSANSADSEY-----KGNLTETINKKTTSDNAVL--LAVKEVAL 49
 DB 424 NAGLTKEKGDGTQIDLNEAATAILNNEKEKQANFTLTKNVNNNALPPDKVARVNAV 483
 QY 50 LSSIDEIAAAGIKKIHQNGGLDTENNHNHNSLLAGAVAI STLKOKLDGKNGE--KEK 107
 DB 484 LETI-----KNNQDTPDIEKSKMEATVAIT-----LNSENLTTPROK 520
 QY 108 IDAAKCSFTTNKLEKHTDLEKGVTDADAKKEAIKAN-STKTGAE--ELGKLFESV 164
 DB 521 QOMLEKAVDVDSFKDTSRAVAIDITG--AVIKNSLTCKKGMILAVGDKNVAS 575
 QY 165 EVLSKAKEMLANSV-----KELTSP-----VAESKPKPSMNSGKDGNT 206
 DB 576 E-LSNAKEKQOLGSLVAKKVEYKIIISPEQOOLMOQNLDKITAEOYTKNDNITEVOGILANP 634
 QY 207 SANSADSE-----VKGPNLTETISKRTES----- 230
 DB 635 AENYTAKTATAIQKVTYKVLDSPTTAIEIKGETLESITIKIVASPLNAVQDKTDIVAGKEAI 694
 QY 231 -----NAVLAVKEVET-----LITS-----IDELAKAI 254
 DB 695 ASHRTMAPTKKIAALESEYETGVAKSTYDLEDKKLTMTGLVDGIYEDKANPEITSEMKAV 754
 QY 255 GKIKNDVSLNDEADHNSLISGAVYLSNLTIKKISAIKSGELAEIEKAKKCSSEFTAK 314
 DB 755 SKGVN-----STAIPEDKQALKPAAS-EAALDRA--TQNFTE 789
 QY 315 KLGKHTDLGKREGVTDNAAKA-----ILKTNNDKTGADE--LEKLFESV 359
 DB 790 GLKGONLDEPKP--RDDIYNKAODIAVALKNVYTVLDANPEKREVESEEVNKKTSILN 847
 QY 360 NLSRAKEMLTN 371
 DB 848 DISKIAIEKVN 859

RESULT 10
 RBP2_P1AVB STANDARD; PRT; 1251 AA.
 ID RBP2_P1AVB
 AC 000799;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)

01-OCT-1996 (Rel. 34, Last annotation update)
DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN RBP2.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9231538; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites."
RL Cell 69:1213-1226(1992).
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M88098; AAA29744.1; -
KW Malaria; Receptor; Membrane.
FT NON_TER 1
FT NON_TER 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 9.0%; Score 166.5; DB 1; Length 1251;
Best Local Similarity 22.6%; Pred. No. 0.95;
Matches 91; Conservative 55; Mismatches 142; Indels 115; Gaps 17;

OY 22 KGNPLTEIKKTKTDSMAVLLAVKEVALLSSIDEIAKAIGKKIHNNGLDPENNINGSI 81
DB 774 KTNELDVHKNIDAKVLEI-----LAHSDI-----DTKQKOSKL 812
OY 82 LA---GAVYISTLIKQKLDGLKNEGLKEKIDAAKCEFTFNKL---KEKHDLGREGYT 135
DB 813 IEMGNQIYIAKVVLLN-----YKNKISSIKSEAVSVAKIGVSKHSHLSKTIIS 863
OY 136 DDADEAT-----LKANGTKGAELGLKFESEVLSKAAKEMLSVSKELTSP 185
DB 864 DKSVDNIILEKQTELQNLNRSFTQKNTNSDSKLEI---KTFESIKNALKTLGE 919
OY 186 VVAESPKKRSMVNNSGKDNSTANSADSEVKGPNLFEISKITTESAAVYLAKEVEITLT 245
DB 920 VNA-----LKASSDNHEVQSKSEPV-NPAUSEIKKEET-----DIDSLMT 959
OY 246 SIDELAK-----AIGKKINDVSLDNEADHNSLISGAVL----- 280
DB 960 ALDELKAKRTCEVSKYKLIKQTVKEISDDELINTIEKNK---ATLAYIKKNVEDT 1015
OY 281 ISNLT--KKAISKDSGLKAEIEAKKCSSEFTAKLGEHTDLCKEGVTDNNAKAIL 338
DB 1016 VODVLTLLNHFMTKQVSNHEPTNFDSKNSSEELTKAVVDSKTIISK-----LKGVII 1068
OY 339 KTN-----NDKTKGADELKFESEVSKLSAAKEM--LNSVSK 374
DB 1069 EVNENMTENTIESSAKETALNELKNNKTKTSLENIYOTSNVKA 1111

RESULT 11
MYS_AEOIR STANDARD; PRT; 1938 AA.
AC P24733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN HEAVY CHAIN, STRIATED MUSCLE.

OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92011595; PubMed=1917970;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Complete primary structure of a scallop striated muscle myosin heavy
RT chain. Sequence comparison with other heavy chains reveals regions
RT that might be critical for regulation."
RL J. Biol. Chem. 266:18469-18476(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=91088319; PubMed=2263488;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
RT muscle myosin heavy chain."
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;
RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA Szent-Gyorgyi A.G., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2.8-A
RT resolution."
RL Nature 368:306-312(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX MEDLINE=96419133; PubMed=8805510;
RA Houdusse A., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2-A
RT resolution: implications for regulation."
RL Structure 4:21-32(1996).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X55714; CAA39247.1; -
DR PIR: S1357; S1357.
DR PIR: A40997; A40997.
DR PDB: 1SCM; 30-APR-94.
DR PDB: 1WDC; 11-JUL-96.
DR InterPro: IPR000048; IO.
DR InterPro: IPR002928; Myosin_head.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF01576; myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.

FT DOMAIN 1 777 MYOSIN HEAD-LIKE.
 FT DOMAIN 778 805 IQ.
 FT DOMAIN 836 1938 ROD-LIKE TAIL (S2 AND LAM DOMAINS).
 FT DOMAIN 836 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 176 183 ATP (BY SIMILARITY).
 FT MOD_RES 693 693 ALKYLATION (SH-1) (BY SIMILARITY).
 FT MOD_RES 703 703 ALKYLATION (SH-2) (BY SIMILARITY).
 FT HELIX 778 821
 FT TURN 822 823
 FT HELIX 825 833
 FT TURN 834 835
 SO SEQUENCE 1938 AA; 222821 MW; ASCCE4127D1A4896 CRC64;

Query Match 8.9%; Score 165.5; DB 1; Length 1938;
 Best Local Similarity 22.3%; Pred. No. 1.7; Mismatches 167; Indels 81; Gaps 12;

Matches 88; Conservative 59; Mismatches 167; Indels 81; Gaps 12;

18 DESVKGNTLEINKKTDNAVLLAVEVALLSIDELAKAIGKIHONGIDT-----73
 DB DEEDAADLGIKKKMEADNA-----NLKKDIGDLENTLOKAEODKAKHNDQISTLOGE 978
 QY 74 ---ENNNGSLAGAVYISTLIKOKDLKNEGLKKEIKIAAKCSEFTTNKIKKEKHTDGL 130
 DB 979 ISODEHIGLKNKEKALEANKKTSIDQAE--EDKCNHLNKLAKKLEQALDELDNDE 1036
 QY 131 KEVYTDADAEAILKANGTYTKGAELGKLFESVEVLSAKAKEMLANVKEITSPVAVES 190
 DB 1037 REKRVKDVDEKA-----KRVKEDOLKTSQEVNEDLEHVKRELEEN-----VRKK 1080
 QY 191 PKRPSWNNNGKDGNTSANSADSVK--GNPLTEISKI--TESNAVLAKEVEITLLISI 247
 DB 1081 EATISSNLSLEDEQNLVLSOLKRIKELOARIELEELAEARNRAKVEKORAEINREL 1140
 QY 248 DELAKAIGK-----KIRNDV---SLDNEADHNGSLISGAVLISNL 284
 DB 1141 EELGERLIDAGCATSAQIETLNKKREAEILKIRIDLEASLQHEAO-----ISAL 1189
 QY 285 IYKKAISAIKDSGLKEIKAKKCCSEFTAKLKGHTDGLKGEVYTDNAKAIILKTNNDK 344
 DB 1190 RKKHODANEMADVOVDOLKVKSLKEDKDKDLKREMDLESO-----MTHMK 1237
 QY 345 TKGADE-LKLFESVYNLSKAKEMLTNSVKELTS 378
 DB 1238 NKGCEKVKKQFESQMSDLNARLEDSQSRINELQS 1272

USOL_YEAST STANDARD; PRT; 1790 AA.

AC P25386;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.
 GN USOL OR INT1 OR YDL058W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCB1_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=91185402; PubMed=2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 RA Yamaseki M.;
 RT "A cytoskeleton-related gene, usol, is required for intracellular
 RT protein transport in Saccharomyces cerevisiae.";
 RL J. Cell Biol. 113:245-260(1991).
 [2]
 RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hostetter M.K., Herman D.J., Bendel C.M., McElellan M., Tao N.,
 RA Kendrick K.E.;

RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Bai Y., Symington L.S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -! FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
 CC COMPLEX.
 CC -! SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
 CC ER AND THE GOLGI COMPLEX.
 CC -! DOMAIN: THE ROD-LIKE TAIL. SEQUENCE IS HIGHLY REPETITIVE. COMPOSED
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
 CC -! SIMILARITY: BELONGS TO THE YDP/USOL/YBL047C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54378; CA38253.1; -;
 DR EMBL; L03188; AAB00143.1; -;
 DR EMBL; U53668; AAB6659.1; -;
 DR PIR; A38455; A38455.
 DR HSSP; P80220; IDIP.
 DR SGD; S0002216; USOL.
 DR InterPro; IPR002017; Spectrin.
 DR TransPort; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
 KW DOMAIN 1 724
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
 FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
 FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
 FT CONFLICT 847 847 G->E (IN REF. 2).
 FT CONFLICT 924 924 E->K (IN REF. 2).
 FT CONFLICT 1253 1253 V->I (IN REF. 2).
 FT CONFLICT 1319 1319 I->V (IN REF. 2).
 FT CONFLICT 1461 1461 N->S (IN REF. 2).
 FT CONFLICT 1581 1581 G->S (IN REF. 2).
 FT CONFLICT 1600 1600 I->V (IN REF. 2).
 FT CONFLICT 1661 1661 R->S (IN REF. 2).
 FT CONFLICT 1772 1772 D->DEPDEE (IN REF. 2).
 SO SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 8.8%; Score 164; DB 1; Length 1790;
 Best Local Similarity 21.6%; Pred. No. 1.8; Mismatches 186; Indels 106; Gaps 17;
 Matches 101; Conservative 75; Mismatches 186; Indels 106; Gaps 17;

3 CNNSGKDG-----NTSANSADSVKGNPLTEINKKTDNAVLLAVEVALLSS 52
 DB 929 CNNLSEKHEIKELVEYKSRFQSHDNLV--AKTEKLKSLANNYKMQA--ENESLIKA 984
 QY 53 IDEIAAKAGKRIHONGID--TENNHNGSLAGAV-----AISTL-----IKQ 95
 DB 985 VEESKNESSTIQLSNQNKTDMSQSEKENQFIRGSIENINIDKTTISDLOQTEELISK 1044
 QY 96 LDKLNKG-----LKEIKDAKKCSEFTTNKKE-----KHYDGLKGEVTD 136
 DB 1045 SSSSDVEYSQISLLEKLETAATVANDENVKISLTTRELEPAELAAVKNLNELETK 1104
 QY 137 ADAKAAIILK-----ANGTKRK-----GAEILGKLFESVLSKAA 171
 DB 1105 LETSKALKAYKENEHLEKEIKQLEKATETKQOULSLRANLSLEKHEHDLAQAQKY 1164
 QY 172 KEMLANVKEITSPV--VAESPKRPSWNNNGKDGNTSANSADSVKGNPLTEISKRT 229
 DB 1165 EEOIANKEROYNEISQINDEITSTQOENESIKKKNDLEGEVAKAMSTBEQSINLKSE 1224
 QY 230 SNAVYLAIVKEY-----ETLLTISIDELAKAIGKIKNDVSLDNEADHNGSLISGAVLISNLI 285


```

DB 1225 IALNQLIKELKKNKTNSLESIKSVSESEVYKIKELODECFNEKESE-----L 1277
QY 286 TKKISAIKSG-----ELKAEIEKAKCSEBFAKLK---GEHTDLCEGVTDNNAKAIL 338
DB 1278 EKLKASEKNSKSYKLELOKESEKIKEELDAKTTELKIQELKINLSKAKESSELSRLK 1337
QY 339 KTNNDKTKADE-LEKL-----FESVKLSKAKEMLNYSKE 375
DB 1338 KTSSEKRNKAEQLEKLKNEIQIKNAQFEKEREKRLNKGSSITTYQYSE 1385

RESULT 13
MYS2_DICDI
ID MYS2_DICDI STANDARD; PRT; 2116 AA.
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
CN MHCA.
CM Dicyostellium discoideum (Slime mold).
OX Eukaryota; Mycetozoa; Dictyostelida; Dictyostellium.
NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from
RN Dicyostellium discoideum."
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN [2]
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN=AX2;
RX MEDLINE=90353583; PubMed=2387408;
RA Lueck-Velometer D., Schleicher M., Grabatin B., Wippler J.,
RT Gerlsch G.;
RN "Replacement of threonine residues by serine and alanine in a
RT phosphorylatable heavy chain fragment of Dicyostellium myosin II."
RL FEBS Lett. 269:239-243(1990).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Magle G., Noegel A., Scheel J., Gerlsch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
RN Dicyostellium myosin heavy chain."
RL FEBS Lett. 227:71-75(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RT Rayment I.;
RN "X-ray structures of the myosin motor domain of Dicyostellium
RT discoidium complexed with MgADP, BeFx and MgADP.ALFA."
RL Biochemistry 34:8960-8972(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RN "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT truncated head of Dictyostellium discoideum myosin to 2.7-A
RT resolution."
RL Biochemistry 34:8973-8981(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RN "X-ray structure of the magnesium(II)-ADP-vanadate complex of the
RT Dicyostellium discoideum myosin motor domain to 1.9-A resolution."
RL Biochemistry 35:5404-5417(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX MEDLINE=97452580; PubMed=9305951;

```

```

RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammaS, and MgAMPNP complexes
RN of the Dicyostellium discoideum myosin motor domain."
RL Biochemistry 36:11619-11628(1997).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RN of Mg.2(3')-O-(N-methylanthranilloyl) nucleotides bound to the
RT Dicyostellium discoideum myosin motor domain."
RL J. Mol. Biol. 274:394-407(1997).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC CORTEX.
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN BE FURTHER
CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHARED
CC SUBFRAGMENT (S2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: DICTYOSTELLUM MYOSIN II HAS NO K(2)DTA ATPASE
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC POSITION (688).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14628; AAA33227.1;
DR PIR: A26655; A26655.
DR PIR: S00250; S00250.
DR PDB: 1MMA; 03-DEC-97.
DR PDB: 1MMD; 17-AUG-96.
DR PDB: 1MNG; 03-DEC-97.
DR PDB: 1MNN; 03-DEC-97.
DR PDB: 1MND; 17-AUG-96.
DR PDB: 1MNE; 17-AUG-96.
DR PDB: 1VOM; 23-DEC-96.
DR PDB: 1LVK; 28-JAN-98.
DR DictyDb: DDO1008; mhca.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ.1.
DR Pfam: PF00663; myosin_head.1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head.1.
DR SMART: SM00015; IQ.1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ.1.
DR KMW: Calmodulin-binding; Actin-binding; 3D-structure;
KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
FT DOMAIN 762 791 IQ.
FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
FT NF_BIND 179 186 ATP.
FT DOMAIN 638 660 ACTIN-BINDING.
FT FT 738 752 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).

```


Matches 84; Conservative 56; Mismatches 147; Indels 98; Gaps 12;

QY 44 KEVALLSSTDEIAAKAIG---KKIQNNGLDTENNHNSLAGAVALSTLQKIDGLK 100
 DB 2208 EEVHQRLRGIEKLRVIEADEKKQHLIAEKLKERERENSL---KDKVENLE 2256
 QY 101 N-----EGLKEKIDAAKCESEFTNKKLEKHTDLDKEGYTADAKVEAIL 144
 DB 2257 RELQNSEENOEVLITDAENSKAEVETLKTQIEEMARSLKVFELD---VTLSEKENLT 2312
 QY 145 KANGTKTGAEBELKLFESVEVL-----SKAKEMLANSVKELTSPVAES 190
 DB 2313 KQIQEKGQGLSELDKLLSSFKSLKEEQAELIQEKESKTAVMQLNQKELMEAV--- 2368
 QY 191 PKKPSVANNSGDGNTSANSADESVPNLTETSKITTESNAVYLAKEVETLLTIDEL 250
 DB 2369 -----AALCGDQELMKATEQSLDPIEEHQRLNRSIEKRLARLEADEKK 2412
 QY 251 ANAIGKKINDVSLDNEADHNSLISGAVYLISNL-----ITKKISAIKDSGELKAEIE 303
 DB 2413 QLCVLDQOLK-----ESEHHDLLKGR--VENLERELEIARTNOEHAALEANSKGEVE 2463
 DB 304 KAKKCESEFTAKLGEHTDL---GREGVTDNNAK-----ATLKTNNDKTKGADEL 351
 DB 2464 TLKAKTEGTMQSLRGLELDVVTIRSEKEDLTNLOKQGERISELEIINSSFENILQKEQ 2523
 QY 352 EKLFESVKNLSSAKAKEMLTNSVKEL 376
 DB 2524 EKV--OMKESKSTAMEMLOTQKEL 2546

RESULT 15
 MSB.CAEEL
 ID MYSB.CAEEL STANDARD: PRT; 1966 AA.
 AC P02366;
 DT 21-JUL-1986 (rel. 01, last sequence update)
 DT 21-JUL-1986 (rel. 01, last sequence update)
 DT 20-AUG-2001 (rel. 40, last annotation update)
 DE MYOSIN HEAVY CHAIN B (MHC B).
 GN UNC-54 OR MYO-4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RP MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 myosin heavy chain gene are not separated by introns.";
 RT Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RL (2)
 RP SEQUENCE OF 850-1966 FROM N.A.
 RP MEDLINE=82272395; PubMed=7202124;
 RA McLeachlan A.D., Karn J.;
 RT "Periodic charge distributions in the myosin rod amino acid sequence
 match cross-bridge spacings in muscle.";
 RT Nature 299:226-231(1982).
 RL (3)
 RP SEQUENCE OF 1876-1966 FROM N.A.
 RP MEDLINE=83232892; PubMed=6571695;
 RA Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
 RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber
 RT nonsense mutations via altered transfer RNA.";
 RL Cell 33:575-583(1983).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
 CC ELEGANS.
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC
 CC EMBL: J01050; AAA28124.1; -;
 CC EMBL: V01494; CAA24738.1; -;
 CC PIR: A02992; MKMW.
 CC DR HSSP: P08799; LMND.
 CC DR InterPro: IPR002928; Myosin_tail.
 CC DR InterPro: IPR001609; myosin_head.
 CC DR Pfam: PF00063; myosin_head.
 CC DR Pfam: PF01576; Myosin_tail; 1.
 CC DR PRINTS: PR00193; MYOSINHEAVY.
 CC DR ProDom: PD000355; myosin_head; 1.
 CC DR SMART: SM00242; MYSC; 1.
 CC KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 CC KW ATP-binding; Methylation; Alkylation; Multigene family.
 CC FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
 CC FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
 CC FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (52).
 CC FT DOMAIN 1165 1176 HINGE.
 CC FT NP_BIND 177 184 ATP (BY SIMILARITY).
 CC FT DOMAIN 665 687 ACTIN-BINDING.
 CC FT DOMAIN 769 783 ACTIN-BINDING.
 CC FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
 CC FT MOD_RES 705 705 ALKYLATION (SH-1).
 CC FT MOD_RES 715 715 ALKYLATION (SH-2).
 CC FT CONFLICT 1337 1337 E -> R (IN REF. 2).
 CC FT CONFLICT 1880 1880 I -> L (IN REF. 2).
 CC SQ SEQUENCE 1966 AA; 225125 MW; B66F0BB2E27B67F CRC64;

Query Match 8.5%; Score 157.5; DB 1; Length 1966;
 Best Local Similarity 23.0%; Pred. No. 4;
 Matches 90; Conservative 69; Mismatches 147; Indels 85; Gaps 18;

QY 16 SADESVAKGNLTETNKKITTSNAVLAVK---EVALLSSTDEL--AAKAIKKRIHQ--NN 69
 DB 1358 SLBEEIEIGKN--ETLRQLSADNADIQQKARFEDEGLKA--DELEDAKRRQAQKINLEQ 1414
 QY 70 GLDTENNHNGL--LAGAVASTLTKOKLDGLKNGLEKIDAAKCESEFTNKKLEKHTD 128
 DB 1415 ALDAANSKNASLEKTSRLVGDLDQAVDVRANGVASALEKKQKGGKIIIDEKRRKTTDD 1474
 QY 129 LKREGVTDADAKAELKANGT---KTKGAEBELKLFESVEVLSSKAKEMLANSVKELTSP 185
 DB 1475 LAAE---LDQAQDLANTSTDLFKAKNAQE--ELAEVVEGLREKNS--LSQEIKDLTDQ 1527
 QY 186 VVAESPKKPSVANNSGDGNTSANSADESVPNLTETSKITTESNAVYLAKEVETLLTIDEL 232
 DB 1528 L-----GEGG---RSVHEKQIIRLEIEKEEQLQHALDEAALEAESK 1569
 QY 233 VYLAKEVETLLSIDLAIAIKKIKN-----DVLSDNEADHNSLISGAVL 280
 DB 1570 VLAQVFEVSOIRSEIERIOKEKEEFENTRKNHRALESQAASLETAKKAKAELR---- 1625

Tues Mar 19 10:57:46 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:46 ; Search time 68.77 Seconds
(without alignments)
418.700 Million cell updates/sec

Title: US-09-596-746a-28

Perfect score: 1859

Sequence: 1 MACNNNSGRKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 378

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	945	50.8	210 2 G70218	outer surface prot
2	938	50.5	2 S69927	outer surface prot
3	879	47.3	210 2 I40144	outer surface prot
4	853	45.9	177 2 I40129	outer surface prot
5	774.5	41.7	211 2 I40145	outer surface prot
6	773.5	41.6	193 2 S70279	outer surface prot
7	700.5	37.7	193 2 S70287	outer surface prot
8	695.5	37.4	211 2 S69918	outer surface prot
9	692.5	37.3	211 2 I40277	outer surface prot
10	690.5	37.1	193 2 S70280	outer surface prot
11	686.5	36.9	211 2 I40268	outer surface prot
12	686.5	36.9	190 2 I40273	outer surface prot
13	685.5	36.9	209 2 I40273	outer surface prot
14	683.5	36.8	191 2 S70278	outer surface prot
15	683.5	36.8	191 2 I40153	outer surface prot
16	675	36.3	194 2 S70277	outer surface prot
17	674.5	36.3	209 2 S69917	outer surface prot
18	670	36.0	212 2 S70254	outer surface prot
19	666.5	35.9	193 2 S70276	outer surface prot
20	664.5	35.7	193 2 S70274	outer surface prot
21	664.5	35.7	211 2 I40278	outer surface prot
22	664	35.7	212 2 S69922	outer surface prot
23	662.5	35.6	191 2 S70284	outer surface prot
24	662.5	35.6	193 2 S70265	outer surface prot
25	662	35.6	194 2 S70268	outer surface prot
26	661.5	35.6	211 2 S69932	outer surface prot
27	657.5	35.4	209 2 I40142	outer surface prot
28	655	35.2	212 2 I40279	outer surface prot
29	650	35.0	212 2 I40143	outer surface prot

30	646	34.7	214 2 S69916	outer surface prot
31	641.5	34.5	193 2 S70286	outer surface prot
32	638	34.3	194 2 S70289	outer surface prot
33	635.5	34.2	191 2 S70288	outer surface prot
34	635.5	34.2	203 2 I40108	outer surface prot
35	631	33.9	212 2 S20543	outer surface prot
36	630.5	33.9	209 2 I40281	outer surface prot
37	630.5	33.9	209 2 I40285	outer surface prot
38	630	33.9	192 2 S70285	outer surface prot
39	628	33.8	210 2 S69925	outer surface prot
40	627	33.7	210 2 S69923	outer surface prot
41	627	33.7	212 2 S69921	outer surface prot
42	622.5	33.5	209 2 S69926	outer surface prot
43	622	33.5	210 2 S69920	outer surface prot
44	619	33.3	210 2 I40272	outer surface prot
45	617.5	33.2	193 2 S70259	outer surface prot

ALIGNMENTS

RESULT 1
G70218
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C:Accession: G70218; I40269; S37726; S70281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; M01D:98065943
A:Accession: G70218
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <KLE>
A:Cross-references: GB:AE000792; NID:93253098; PIDN:AAC66329.1; PID:92689901; TIGR:BB
A:Experimental source: strain B31
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: I40269; M01D:96025162
A:Accession: I40269
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:D49497; NID:9707092; PIDN:BA08457.1; PID:9769684
R:Daunis-Heipke, S.; Fuchs, R.; Mottz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os
A:Reference number: S37726; M01D:93268136
A:Accession: S37726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <JAU>
A:Cross-references: EMBL:X69596; NID:9311391; PIDN:CAA49306.1; PID:9311392
R:Liver, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; M01D:9626448
A:Accession: S70281
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210 <LIV>
A:Cross-references: EMBL:I42887; NID:9858715; PIDN:AAB36995.1; PID:91695212
A:Experimental source: strain Ip2
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match
 Best Local Similarity 50.8%; Score 945; DB 2; Length 210;
 Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNPLTEINKITDSNAVLAVKVEALLSSIDEIAKA 60
 Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLAVKVEALLSSIDEIAKA 60
 QY 61 IGRKHQNNGLDTENNHNSLAGAYASTLTKOKLDLKNKGKKEKIDAAKCKSEFTFN 76
 Db 77 IGRKHQNNGLDTENNHNSLAGAYASTLTKOKLDLKNKGKKEKIDAAKCKSEFTFN 120
 QY 121 KLEKEHTDGLGKGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSYK 180
 Db 137 KLEKEHTDGLGKGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSYK 196
 QY 181 ELTSPVAESPKRP 194
 Db 197 ELTSPVAESPKRP 210

RESULT 2
 S69927
 outer surface protein C precursor - Lyme disease spirochete (strain PKA)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 A:Variety: strain PKA
 C:Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
 C:Accession: S69927; S72669
 R:Jauris-Helpe, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.
 J. Clin. Microbiol. 33: 1860-1866, 1995
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
 A:Reference number: I40047; MUID:95395018
 A:Accession: S69927
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-210 <JAU>
 A:Cross-references: EMBL:X69589
 R:Jauris, S.
 A:Experimental source: strain PKA
 A:Submitted to the EMBL Data Library, February 1994
 A:Reference number: S72669
 A:Accession: S72669
 A:Molecule type: DNA
 A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAW>
 A:Cross-references: EMBL:X69589
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match
 Best Local Similarity 50.5%; Score 938; DB 2; Length 210;
 Matches 188; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNPLTEINKITDSNAVLAVKVEALLSSIDEIAKA 60
 Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKITDSNAVLAVKVEALLSSIDEIAKA 60
 QY 61 IGRKHQNNGLDTENNHNSLAGAYASTLTKOKLDLKNKGKKEKIDAAKCKSEFTFN 76
 Db 77 IGRKHQNNGLDTENNHNSLAGAYASTLTKOKLDLKNKGKKEKIDAAKCKSEFTFN 120
 QY 121 KLEKEHTDGLGKGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSYK 180
 Db 137 KLEKEHTDGLGKGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSYK 196
 QY 181 ELTSPVAESPKRP 194
 Db 197 ELTSPVAESPKRP 210

outer surface protein C precursor - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
 C:Accession: I40144; S70282
 R:Stevenson, B.; Barthold, S.W.
 FEMS Microbiol. Lett. 124, 367-372, 1994
 A:Title: Expression and sequence of outer surface protein C among North American isol
 A:Reference number: I40144; MUID:95154673
 A:Accession: I40144
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-210 <RES>
 A:Cross-references: EMBL:U04281; NID:9434663; PIDN:AA37001.1; PID:9434664
 R:Livley, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70282
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 19-210 <LIV>
 A:Cross-references: EMBL:L42893; NID:958721; PIDN:AA37001.1; PID:91695218
 C:Experimental source: strain 297
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match
 Best Local Similarity 47.3%; Score 879; DB 2; Length 210;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 198 NSNGKDGNTSANSADSVKGNPLTEISKITDSNAVLAVKVEALLSSIDEIAKA 257
 Db 20 NSNGKDGNTSANSADSVKGNPLTEISKITDSNAVLAVKVEALLSSIDEIAKA 79
 QY 258 IKNVSLDNEADHNGSLISGAYLISLITKISAINDSGELAEIEKAKCKSEFTATLK 317
 Db 80 IKNVSLDNEADHNGSLISGAYLISLITKISAINDSGELAEIEKAKCKSEFTATLK 139
 QY 318 GEHTDLGKGVTDADAKKAILKANGTKTKGAELGKLFESVYNSKAKEMLINSVELT 377
 Db 140 GEHTDLGKGVTDADAKKAILKANGTKTKGAELGKLFESVYNSKAKEMLINSVELT 199
 QY 378 S 378
 Db 200 S 200

RESULT 4
 I40129
 outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
 C:Accession: I40129; S54199
 R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
 J. Bacteriol. 177, 3036-3044, 1995
 A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
 A:Reference number: I40104; MUID:95286481
 A:Accession: I40129
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-177 <RES>
 A:Cross-references: EMBL:X84783; NID:9793825; PIDN:CAA59254.1; PID:9793826
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match
 Best Local Similarity 45.9%; Score 853; DB 2; Length 177;
 Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

RESULT 3
 I40144

OY 11 NTSANSADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAKIGKIHONG 70
 DB 1 NTSANSADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAKIGKIHONG 60
 OY 71 LDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFNKLEKHTDGL 130
 DB 61 LDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFNKLEKHTDGL 120
 OY 131 KCGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVKELTSPV 187
 DB 121 KCGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVKELTSPV 177

RESULT 5
 I40145
 outer surface protein C precursor - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
 C:Accession: I40145
 R:Stevenson, B.; Barthold, S.W.
 PMS Microbiol. Lett. 124, 367-372, 1994
 Title: Expression and sequence of outer surface protein C among North American isolate
 Reference number: I40143; MUID:95154673
 A:Accession: I40145
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-211 <RES>
 A:Cross-References: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.7%; Score 774.5; DB 2; Length 211;
 Best Local Similarity 81.0%; Pred. No. 2.7e-29;
 Matches 158; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

OY 1 MACNNSGKDCN-TSANSADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAK 59
 DB 17 ISCNNSGKDCNASTNPADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAK 76
 OY 60 AIGKRIHONGGLDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFN 119
 DB 77 AIGKRIHONGGLDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFN 136
 OY 120 NKLKEHTDLGKEGYTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSV 179
 DB 137 NKLKSHHTLGGQDQDDAKKAILRTHTNTKDKGAELDKLFRVAVNLKAAKEMLANSV 196

180 KELTSPVAESPKP 194
 197 KELTSPVAESPKP 211

RESULT 6
 S70279
 outer surface protein C - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70279
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70279
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <LIV>
 A:Cross-References: EMBL:I42898; NID:g858723; PIDN:AAB37007.1; PID:g1695222
 A:Experimental source: strain 25015
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.6%; Score 773.5; DB 2; Length 193;
 Best Local Similarity 81.9%; Pred. No. 2.7e-29;
 Matches 158; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

OY 3 CNNSGKDCN-TSANSADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAKI 61
 DB 1 CNNSGKDCNASTNPADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAKI 60
 OY 62 GKRIHONGGLDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFNK 121
 DB 61 GKRIHONGGLDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFNK 120
 OY 122 LKEKHTDLGKEGYTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSV 181
 DB 121 LKSHHTLGGQDQDDAKKAILRTHTNTKDKGAELDKLFRVAVNLKAAKEMLANSV 180
 OY 182 LTPSPVAESPKP 194
 DB 181 LTPSPVAESPKP 193

RESULT 7
 S70287
 outer surface protein C - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70287
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70287
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <LIV>
 A:Cross-References: EMBL:I42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
 A:Experimental source: strain 28354
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.7%; Score 700.5; DB 2; Length 193;
 Best Local Similarity 75.1%; Pred. No. 5.9e-26;
 Matches 145; Conservative 17; Mismatches 30; Indels 1; Gaps 1;

OY 3 CNNSGKDCN-TSANSADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAKI 62
 DB 1 CNNSGKDCNASTNPADSVKGNPLTEINKKITDSNAVLAVKEVEALLSIDEIAAKI 60
 OY 63 KRIHONGGLDTENNHNHSLAGAVAI STLTKOKLDGLKNGEKLKEKIDAAKCSFTFNK 121
 DB 61 KRIHONGGLAVENGHTLAGAVITSTLTKOKLDGLKNGEKLKEKIDAAKCSFTFNK 120
 OY 122 LKEKHTDLGKEGYTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSV 181
 DB 121 LKEHHTLGGQDQDDAKKAILRTHTNTKDKGAELDKLFRVAVNLKAAKEMLANSV 180
 OY 182 LTPSPVAESPKP 194
 DB 181 LTPSPVAESPKP 193

RESULT 8
 S69918
 outer surface protein C precursor - Lyme disease spirochete (strain p8e)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 A:Variety: strain p8e
 C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
 C:Accession: S69918; S72674; I40103
 R:tautis-Helpke, S.; Diegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek

J. Clin. Microbiol. 33, 1860-1866, 1995
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia burgdorferi*
 A:Reference number: 140047, MUID:9595018
 A:Accession: S69918
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-211 <CDS>
 A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022
 A:Experimental source: strain pRe
 R:Roessler, D.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: S72674
 A:Accession: S72674
 A:Molecule type: DNA
 A:Residues: 1-152, 'E', 154-211 <RES>
 A:Cross-references: EMBL:X81522; NID:9872021; PIDN:CAA57242.1; PID:9872022
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.4%; Score 695.5; DB 2; Length 211;
 Best Local Similarity 77.8%; Pred. No. 1.1e-25;
 Matches 152; Conservative 12; Mismatches 29; Indels 3; Gaps 3;

1 MACNSGKDGNTSANSADSVKGNLTETNKRTDSNAVLAVKEVEALLSSIDEIAKA 60
 : :::
 17 ISCNNSGKDGNTSANSADSVKGNLTETNKRTDSNAVLAVKEVEALLSSIDEIAKA 60
 : :::
 61 IGRKHQNNGLDTENNHNSLLAGAVASTLKOKLDGLK-NEGKREIDAKKCEFT 119
 : :::
 76 IGRKHQNNGLDTENNHNSLLAGAVASTLKOKLDGLK-NEGKREIDAKKCEFT 119
 : :::
 120 NKLEKHTDLGKEGYTDADAEALIKANGT-KTGAELGLFESEVLSKAKEMLANS 178
 : :::
 136 TKLKHQAOLGIGYTDNNAKAILKANAKGDKGVELEKLSGLSELSKAKEMLANS 195
 : :::
 179 VKELTSPVAESPKRP 194
 : :::
 196 VKELTSPVAESPKRP 211
 : :::

RESULT 9
 140277
 outer surface protein C precursor - *Borrelia afzelii*
 C:Species: *Borrelia afzelii*
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
 C:Accession: 140277
 R:Fukunaga, M.; Hamase, A.
 J. Clin. Microbiol. 33, 2415-2420, 1995
 A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato
 A:Reference number: 140269; MUID:96025162
 A:Accession: 140277
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-211 <RES>
 A:Cross-references: GB:DA9501; NID:9707096; PIDN:BA08461.1; PID:9769688
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.3%; Score 692.5; DB 2; Length 211;
 Best Local Similarity 74.5%; Pred. No. 1.1e-25;
 Matches 146; Conservative 20; Mismatches 27; Indels 3; Gaps 3;

1 MACNSGKDGNTSANSADSVKGNLTETNKRTDSNAVLAVKEVEALLSSIDEIAKA 59
 : :::
 17 ISCNNSGKDGNTSANSADSVKGNLTETNKRTDSNAVLAVKEVEALLSSIDEIAKA 59
 : :::
 60 AIGKHHQNNGLDTENNHNSLLAGAVASTLKOKLDGLK-EGKREIDAKKCEFT 118
 : :::
 77 AIGKHHQNNGLDTENNHNSLLAGAVASTLKOKLDGLK-EGKREIDAKKCEFT 118
 : :::
 119 TNKLEKHTDLGEGYTDADAEALIKANGT-KTGAELGLFESEVLSKAKEMLANS 178
 : :::

136 TKLKHQAOLGIGYTDNNAKAILKANAKGDKGVELEKLSGLSELSKAKEMLANS 195
 : :::
 179 VKELTSPVAESPKRP 194
 : :::
 196 VKELTSPVAESPKRP 211
 : :::

RESULT 10
 S70280
 outer surface protein C - Lyme disease spirochete
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70280
 R:Livey, J.; Gibbs, C.P.; Schuster, R.; Dörner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in *OspC* variation in Lyme disease
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70280
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <LIV>
 A:Cross-references: EMBL:L42868; NID:9858735; PIDN:AA037011.1; PID:91695226
 A:Experimental source: strain Z57
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 37.1%; Score 690.5; DB 2; Length 193;
 Best Local Similarity 77.8%; Pred. No. 1.7e-25;
 Matches 151; Conservative 10; Mismatches 30; Indels 3; Gaps 3;

3 CNSGKDGNTSANSADSVKGNLTETNKRTDSNAVLAVKEVEALLSSIDEIAKAIG 62
 : :::
 1 CNSGKDGNTSANSADSVKGNLTETNKRTDSNAVLAVKEVEALLSSIDEIAKAIG 62
 : :::
 63 KKHQNNGLDTENNHNSLLAGAVASTLKOKLDGLK-NEGKREIDAKKCEFT 121
 : :::
 60 KKHQNNGLDTENNHNSLLAGAVASTLKOKLDGLK-NEGKREIDAKKCEFT 121
 : :::
 122 LKREKHTDLGKEGYTDADAEALIKANGT-KTGAELGLFESEVLSKAKEMLANS 180
 : :::
 120 LKDNHQAOLGIGYTDNNAKAILKANAKGDKGVELEKLSGLSELSKAKEMLANS 179
 : :::
 181 ELTSPVAESPKRP 194
 : :::
 180 ELTSPVAESPKRP 193
 : :::

RESULT 11
 140268
 outer surface protein C precursor - Lyme disease spirochete
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
 C:Accession: 140268
 R:Margolis, N.; Hogan, D.; Cleplak, W.
 Gene 143, 105-110, 1994
 A:Title: Homology between *Borrelia burgdorferi* *OspC* and members of the family of *Borrelia burgdorferi*
 A:Reference number: 140268; MUID:94259285
 A:Accession: 140268
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-211 <RES>
 A:Cross-references: GB:L25413; NID:9495735; PIDN:AAA22956.1; PID:9495736
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.9%; Score 686.5; DB 2; Length 211;
 Best Local Similarity 73.8%; Pred. No. 2.9e-25;
 Matches 144; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 60
 DB 17 ISCNNSGKGDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 76
 OY 61 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNK-EGLEKIDAAKKSEFTTK 119
 DB 77 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNK-EGLEKIDAAKKSEFTTK 136
 OY 120 NKLEKHTDLGREGVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 179
 DB 137 KKLDSNADLGKHNATDADSKAAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 196
 OY 180 KELTSPVVAESPCKP 194
 DB 197 KELTSPVVAESPCKP 211

RESULT 12

outer surface protein C - Lyme disease spirochete
 Species: Borrelia burgdorferi (Lyme disease spirochete)
 Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70273
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70273
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-190 <LIV>
 A:Cross-References: EMBL:L42870; NID:9858737; PIDN:AAB37013.1; PID:g1695228
 A:Experimental source: strain VSDA
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.9%; Score 686; DB 2; Length 190;
 Best Local Similarity 75.6%; Pred. No. 2.7e-25;
 Matches 146; Conservative 19; Mismatches 24; Indels 4; Gaps 3;

OY 3 CANSNGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 62
 DB 1 CANSNG--GDTASTNPDESAGKPNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 57
 OY 63 KKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKN-NEGLEKIDAAKKSEFTTK 121
 DB 58 KKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKN-NEGLEKIDAAKKSEFTTK 117
 OY 122 LKREKHTDLGREGVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 181
 DB 118 LKREKHTDLGREGVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 177
 OY 182 LTPSPVVAESPCKP 194
 DB 178 LTPSPVVAESPCKP 190

RESULT 13

outer surface protein C precursor - Borrelia afzelii
 C:Species: Borrelia afzelii
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
 C:Accession: I40273
 R:Fukunaga, M.; Hamase, A.
 J. Clin. Microbiol. 33, 2415-2420, 1995
 A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
 A:Reference number: I40269; MUID:96025162
 A:Accession: I40273
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-209 <RES>
 A:Cross-References: GB:D49379; NID:q1041107; PIDN:BAA08377.1; PID:g1041108
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.9%; Score 685.5; DB 2; Length 209;
 Best Local Similarity 74.4%; Pred. No. 3.2e-25;
 Matches 145; Conservative 21; Mismatches 26; Indels 3; Gaps 3;

OY 1 MACNSGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 60
 DB 17 ISCNNSGKGDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 75
 OY 61 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKN-EGLEKIDAAKKSEFTTK 119
 DB 76 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKN-EGLEKIDAAKKSEFTTK 134
 OY 120 NKLEKHTDLGREGVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 179
 DB 135 KKLDSNADLGKHNATDADSKAAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 194
 OY 180 KELTSPVVAESPCKP 194
 DB 195 KELTSPVVAESPCKP 209

RESULT 14

outer surface protein C - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70278
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70278
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-191 <LIV>
 A:Cross-References: EMBL:L42871; NID:9858738; PIDN:AAB37014.1; PID:g1695229
 A:Experimental source: strain VSD461
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.8%; Score 683.5; DB 2; Length 191;
 Best Local Similarity 75.1%; Pred. No. 3.5e-25;
 Matches 145; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

OY 3 CANSNGKDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 62
 DB 1 CANSNGKGDGNTSANSADSVKGNLTETINKKTTSDSNVAVLAVKEVEALLSSIDEIAAKA 59
 OY 63 KKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKN-EGLEKIDAAKKSEFTTK 121
 DB 60 KKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKN-EGLEKIDAAKKSEFTTK 118
 OY 122 LKREKHTDLGREGVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 181
 DB 119 LKREKHTDLGREGVTDADAKKAILKANGTKTGAEELKLFESVEVLSKAAREMLANSV 178
 OY 182 LTPSPVVAESPCKP 194
 DB 179 LTPSPVVAESPCKP 191

RESULT 15

outer surface protein C - Lyme disease spirochete (fragment)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000

C:Accession: I40153
R:Fung, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
Infect. Immun. 62, 3213-3221, 1994
A:Title: Humoral immune response to outer surface protein C of *Borrelia burgdorferi* in I
A:Reference number: I40153; MUID:94314437
A:Accession: I40153
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <RMS>
A:Cross-references: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:53 ; Search time 124.19 Seconds
(without alignments)
445.212 Million cell updates/sec

Title: US-09-596-746A-28

Perfect score: 1859

Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 378

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Query Match	Length	DB	ID	Description
1	943	50.7	192	2	O9S3P3 borrelia bu
2	927	49.9	191	2	P70818 borrelia bu
3	921	49.5	192	2	O9R7B1 borrelia bu
4	908	48.8	193	2	O9R853 borrelia bu
5	893	48.0	200	2	O9R0R9 borrelia bu
6	885	47.6	182	2	O9R7B2 borrelia bu
7	885	47.6	191	2	O31120 borrelia bu
8	879	47.3	192	2	O9S3P2 borrelia bu
9	879	47.3	210	2	O44719 borrelia bu
10	853	45.9	177	2	O44999 borrelia bu
11	843.5	44.6	178	2	O9REH7 borrelia bu
12	829	44.6	175	2	O9R7B4 borrelia bu
13	802	43.1	163	2	O9R7A9 borrelia bu
14	786	42.3	165	2	O9AGB1 borrelia bu
15	774.5	41.7	211	2	O44720 borrelia bu
16	773.5	41.6	193	2	P94237 borrelia bu
17	773	38.0	159	2	O9R7B0 borrelia bu
18	707	38.0	193	2	O31115 borrelia bu
19	700.5	37.7	193	2	P94234 borrelia bu

20	696.5	37.5	211	2	O44977 borrelia bu
21	692.5	37.3	211	2	O49576 borrelia af
22	690.5	37.1	193	2	P94242 borrelia bu
23	686.5	36.9	211	2	O44976 borrelia bu
24	686	36.9	190	2	P94244 borrelia bu
25	686	36.9	212	2	O9KIMS borrelia af
26	685.5	36.9	209	2	O44671 borrelia af
27	683.5	36.8	191	2	O44726 borrelia bu
28	683.5	36.8	191	2	P94245 borrelia bu
29	681	36.6	212	2	O44705 borrelia bu
30	675	36.3	194	2	P94247 borrelia bu
31	670.5	36.1	194	2	O31122 borrelia bu
32	670	36.0	212	2	O44727 borrelia bu
33	667	35.9	201	2	P96571 borrelia ja
34	666.5	35.9	193	2	P94231 borrelia bu
35	666.5	35.9	211	2	O926C7 borrelia va
36	666	35.8	201	2	P96573 borrelia ja
37	664.5	35.7	193	2	P94233 borrelia bu
38	664.5	35.7	202	2	P96508 borrelia af
39	664.5	35.7	209	2	O44883 borrelia bu
40	664.5	35.7	211	2	O49577 borrelia af
41	664	35.7	203	2	O50624 borrelia af
42	664	35.7	212	2	O44670 borrelia af
43	662.5	35.6	191	2	O9S3P0 borrelia bu
44	662.5	35.6	193	2	P94226 borrelia bu
45	662	35.6	194	2	P94229 borrelia bu

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	192 AA.
O9S3P3	O9S3P3			
AC	O9S3P3			
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)			
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)			
DT	01-JUN-2001 (TREMREL. 17, Last annotation update)			
DE	OUTER SURFACE PROTEIN C (FRAGMENT).			
GN	OSPC.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID-139;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-IP2;			
RX	MEDLINE-96296448; PubMed-8709845;			
RA	Lively I., Gibbs C.P., Schuster R., Dorner F.;			
RT	"Evidence for lateral transfer and recombination in OspC variation in			
RL	Lyme disease Borrelia."			
RM	Mol. Microbiol. 18:257-269(1995).			
DR	EMBL: LA2887; AAB36995.1; "			
DR	InterPro: IPR001800; Lipoprotein_6.			
DR	Pfam: PF01441; Lipoprotein_6; 1.			
DR	ProDom: PD001149; Lipoprotein_6; 1.			
FT	NON_TER			
FT	NON_TER			
SO	SEQUENCE			
Query Match	50.7%;	Score 943;	DB 2;	Length 192;
Best Local Similarity	99.08;	Pred. No. 2.2e-36;		
Matches 190;	Conservative	1;	Mismatches	1;
			Indels	0;
			Gaps	0;
OY	3	CNNSGKDGNTSANSADSVKGNLTETINKITDSENAVLAVKEVEALLSIDETAAKIG 62		
Db	1	CNNSGKDGNTSANSADSVKGNLTETSKITDSENAVLAVKEVEALLSIDETAAKIG 60		
OY	63	KTIHNNGLDTENNHNSLLAGAVAIITLTKOKIDGLKNEGLKEDIKAACKSTFTNNKL 122		
Db	61	KTIHNNGLDTENNHNSLLAGAVAIITLTKOKIDGLKNEGLKEDIKAACKSTFTNNKL 120		
OY	123	KEKHTDGLKGVYDADAKKAILKANGTKTGAEELGLFESVEVLSKAANKEMLTNSVKEL 182		

Db 121 KEKHTDLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180
QY 183 TSPVVAESPKKP 194
181 TSPVVAESPKKP 192

RESULT 2

ID P70818 PRELIMINARY; PRT; 191 AA.

DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE OUTER SURFACE PROTEIN (FRAGMENT).

GN OSPC.
OR Borrelia burgdorferi (Lyme disease spirochete).
OX Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RP SEQUENCE FROM N.A.
RC STRAIN=2-1498 CA4;

RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; L81131; AAB06569.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1

SO SEQUENCE 191 AA; 20126 MW; D2B9B1C82B4DC3C0 CRC64;

Query Match 49.9%; Score 927; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 1.2e-35;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 NNSGKDGNTSANSADSVKGPMLTEINKRTTDSNAVLAVKEVEALLSSIDEIAAKA 63
Db 1 NNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSSIDEIAAKA 60
QY 64 KIHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNEGKLEKIDAKKCEFTFN 123
Db 61 KIHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNEGKLEKIDAKKCEFTFN 120
QY 124 EKHTDGLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKELT 183
121 EKHTDGLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVKELT 180
QY 184 SPVVAESPKKP 194
181 SPVVAESPKKP 191

RESULT 3

ID Q9R7B1 PRELIMINARY; PRT; 192 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OR Borrelia burgdorferi (Lyme disease spirochete).
OX Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RP SEQUENCE FROM N.A.
RC STRAIN=TESTS;

RX MEDLINE=97478003; PubMed=9336916;

RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997)."

RN [2]
RP SEQUENCE FROM N.A.

RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91798; AAB81895.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1

SO SEQUENCE 192 AA; 20297 MW; 6770502A20A6764 CRC64;

Query Match 49.5%; Score 921; DB 2; Length 192;
Best Local Similarity 97.9%; Pred. No. 2.2e-35;
Matches 186; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTNSANSADSVKGPMLTEINKRTTDSNAVLAVKEVEALLSSIDEIAKA 60
Db 3 ISCNNSGKDGTNSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSSIDEIAKA 62
QY 61 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNEGKLEKIDAKKCEFTFN 120
Db 63 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNEGKLEKIDAKKCEFTFN 122
QY 121 KIKKHTDGLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVK 180
Db 123 KIKKHTDGLGKRGVTDADAKKAIILKNTGKTGAEELGKLFESVYLSKAKEMLANSVK 182
QY 181 ELTSPVVAES 190
183 ELTSPVVAES 192

RESULT 4

ID Q9RR53 PRELIMINARY; PRT; 193 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OR Borrelia burgdorferi (Lyme disease spirochete).
OX Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;

RP SEQUENCE FROM N.A.
RC STRAIN=OC1;

RX MEDLINE=99091544; PubMed=9872945;
RA Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT "Genetic diversity of ospC in a local population of Borrelia
RL burgdorferi sensu stricto.";
DR EMBL; AF029860; AAB86543.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1

SO SEQUENCE 193 AA; 20502 MW; 5EFDB5AF8986D1E CRC64;

Query Match 48.8%; Score 908; DB 2; Length 193;
Best Local Similarity 97.9%; Pred. No. 8.7e-35;
Matches 183; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTNSANSADSVKGPMLTEINKRTTDSNAVLAVKEVEALLSSIDEIAKA 60
Db 7 ISCNNSGKDGTNSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSSIDEIAKA 66
QY 61 ICKKHONNGLDTEENNHNHSLAGAYASTLIKOKLDGLKNEGKLEKIDAKKCEFTFN 120

```

DB 67 IGGKIHONNCLDTEENNHNGLAGAVAI STLIRKOKLDGLKNEGIKEKIDAKKCEFTFTN 126
QY 121 KLEKHTDGLKEGVTADADAEALITKANGTKTGAEBELGKLFESVEVLSKAKEMLANSVK 180
DB 127 KLEKHTDGLKEGVTADADAEALITKNGTKTGAEBELGKLFESVEVLSKAKEMLANSVK 186
QY 181 ELTSPV 187
DB 187 ELTSPV 193

RESULT 5
Q9ROR9 PRELIMINARY; PRT; 200 AA.
ID Q9ROR9;
AC Q9ROR9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
EN OSPC.
NC Borrelia burgdorferi (Lyme disease spirochete).
NU plasmid cp26.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=27;
RX MEDLINE=20002545; PubMed=10531219;
RA Hofmeister E.K., Glas G.E., Childs J.E., Persing D.H.;
RT "Population dynamics of a naturally occurring heterogeneous mixture of
  Borrelia burgdorferi clones."
RL Infect. Immun. 67:5709-5716(1999).
DR EMBL: AF074464; AAD23911.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KW plasmid.
FT NON_TER
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BB6B6B CRC64;

Query Match 48.0%; Score 893; DB 2; Length 200;
Best Local Similarity 97.8%; Pred. No. 4,4e-34;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTANSANSADESVKGNLTLEINKKITDSNAVLAVKEVALLSIDEIATAKA 60
DB 17 ICSNNSGKDGNTANSANSADESVKGNLTLEISKKITDSNAVLAVKEVALLSIDEIATAKA 76
DB 61 IGGKIHONNCLDTEENNHNGLAGAVAI STLIRKOKLDGLKNEGIKEKIDAKKCEFTFTN 120
DB 77 IGGKIHONNCLDTEENNHNGLAGAVAI STLIRKOKLDGLKNEGIKEKIDAKKCEFTFTN 136
QY 121 KLEKHTDGLKEGVTADADAEALITKANGTKTGAEBELGKLFESVEVLSKAKEMLANSVK 180
DB 137 KLEKHTDGLKEGVTADADAEALITKNGTKTGAEBELGKLFESVEVLSKAKEMLANSVK 196
QY 181 ELTS 184
DB 197 ELTS 200

RESULT 6
Q9R7B2 PRELIMINARY; PRT; 182 AA.
ID Q9R7B2;
AC Q9R7B2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

```

```

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2ETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
  U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2ETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91797; AAB81894.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match 47.6%; Score 885; DB 2; Length 182;
Best Local Similarity 98.9%; Pred. No. 9,1e-34;
Matches 180; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 GKDGTANSANSADESVKGNLTLEINKKITDSNAVLAVKEVALLSIDEIATAKAIGKKIH 66
DB 1 GKGNTANSANSADESVKGNLTLEISKKITDSNAVLAVKEVALLSIDEIATAKAIGKKIH 60
QY 67 QNNCLDTEENNHNGLAGAVAI STLIRKOKLDGLKNEGIKEKIDAKKCEFTFTN 126
DB 61 QNNCLDTEENNHNGLAGAVAI STLIRKOKLDGLKNEGIKEKIDAKKCEFTFTN 120
QY 127 TDGLKEGVTADADAEALITKANGTKTGAEBELGKLFESVEVLSKAKEMLANSVELTSPV 186
DB 121 TDGLKEGVTADADAEALITKNGTKTGAEBELGKLFESVEVLSKAKEMLANSVELTSPV 180
QY 187 VA 188
DB 181 VA 182

RESULT 7
O31120 PRELIMINARY; PRT; 191 AA.
ID O31120;
AC O31120;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OC10;
RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF029869; AAB86552.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 191 AA; 20502 MW; E4FB56BCB61740F9 CRC64;

Query Match 47.6%; Score 885; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 9,6e-34;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 198 NNSGKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 257
 DB 10 NNSGKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 69
 QY 258 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCESEFTAKLK 317
 DB 70 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCESEFTAKLK 129
 QY 318 GEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAAREMLTNSVKELT 377
 DB 130 GEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAAREMLTNSVKELT 189
 QY 378 S 378
 DB 190 S 190

LT 8
 O953P2
 ID O953P2 PRELIMINARY: PRT: 192 AA.
 AC O953P2:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RX MEDLINE=96296448; PubMed=8709845;
 RA Livey I., Gibbs C.P., Schuster R., Dorner F.,
 RT "Evidence for lateral transfer and recombination in OSpC variation in
 RL Lyme disease Borrelia.";
 DR Mol. Microbiol. 18:257-269(1995).
 DR EMBL: L42893; AAB37001.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 192
 SQ SEQUENCE 192 AA: 20472 MW: 46AC8F93E4DFED6C CRC64;

Query Match 47.3%; Score 879; DB 2; Length 192;
 Best Local Similarity 99.4%; Pred. No. 1,8e-33;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 198 NNSGKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 257
 DB 10 NNSGKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 69
 QY 258 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCESEFTAKLK 317
 DB 70 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCESEFTAKLK 129
 QY 318 GEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAAREMLTNSVKELT 377
 DB 130 GEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAAREMLTNSVKELT 189
 QY 378 S 378
 DB 182 S 182

RESULT 9
 ID O44719 PRELIMINARY: PRT: 210 AA.
 AC O44719:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OSpC.
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB19;
 RX MEDLINE=95154673; PubMed=7851744;
 RA Stevenson B., Barthold S.W.;
 RT "Expression and sequence of outer surface protein C among North
 RL American isolates of Borrelia burgdorferi.";
 DR EMBL: U04281; AAC43297.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 210 AA: 22526 MW: 2D672991D584E4EB CRC64;

Query Match 47.3%; Score 879; DB 2; Length 210;
 Best Local Similarity 99.4%; Pred. No. 2e-33;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 198 NNSGKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 257
 DB 20 NNSGKGNTSANSADSVKGPNTLEISKRTESNAVLAKEVEETLTSIDELAKAIGKK 79
 QY 258 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCESEFTAKLK 317
 DB 80 IKNDVSLDNEADHNGSLISGAYLISNLTITKRTISAIKDNGSGELKAEIEKAKKCESEFTAKLK 139
 QY 318 GEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAAREMLTNSVKELT 377
 DB 140 GEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESVKNLSKAAREMLTNSVKELT 199
 QY 378 S 378
 DB 200 S 200

RESULT 10
 ID O44999 PRELIMINARY: PRT: 177 AA.
 AC O44999:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TJXW;
 RX MEDLINE=95286481; PubMed=7768799;
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 RT "Evolution of the Borrelia burgdorferi outer surface protein OSpC.";
 DR EMBL: X84783; CAAS9234.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 177
 SQ SEQUENCE 177 AA: 18800 MW: 156671B9614E7A2D CRC64;

Query Match 45.9%; Score 853; DB 2; Length 177;

Best Local Similarity 98.3%; Pred. No. 2,5e-32;
Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNNG 70
|||||
DB 1 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNNG 60
|||||

QY 71 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAAKCSFTFNKLEKHTDIDG 130
|||||
DB 61 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAAKCSFTFNKLEKHTDIDG 120
|||||

QY 131 KEGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAEMLANSKELTSPVY 187
| |||||
DB 121 KEGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAEMLANSKELTSPVY 177
| |||||

RESULT 11
Q9REH7 PRELIMINARY: PRT: 178 AA.

AC Q9REH7 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
RT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B31.
RX MEDLINE=94075528; PubMed=8253951;
RA Theisen M., Frederixsen B., Ledech A.M., Vuust J., Hansen K.;
RT "Polymorphism in ospC gene of Borrelia burgdorferi and
RT immunoreactivity of OspC protein: implications for taxonomy and for
RT use of OspC protein as a diagnostic antigen.";
RL J. Clin. Microbiol. 31:2570-2576(1993).
DR EMBL: X73622; CAA52001.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1 178
FT SEQUENCE 178 AA; 18894 MW; D619AEC646F14EB CRC64;

Query Match 45.4%; Score 843.5; DB 2; Length 178;
Best Local Similarity 97.8%; Pred. No. 6,8e-32;
Matches 174; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

DB 11 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNNG 70
|||||
DB 1 NTSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAAKAIGKKIHQNNG 60
|||||

QY 71 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAAKCSFTFNKLEKHTDIDG 129
|||||
DB 61 LPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAAKCSFTFNKLEKHTDIDG 120
|||||

QY 130 KEGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAEMLANSKELTSPVY 187
| |||||
DB 121 KEGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKAEMLANSKELTSPVY 178
| |||||

RESULT 12
Q9R7B4 PRELIMINARY: PRT: 175 AA.

AC Q9R7B4 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
RT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT U.S.A.'";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91792; AAB81889.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1 175
FT SEQUENCE 175 AA; 18573 MW; EC059E7BD3AC3250 CRC64;

Query Match 44.6%; Score 829; DB 2; Length 175;
Best Local Similarity 97.6%; Pred. No. 3,1e-31;
Matches 166; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAKA 60
| : |||||
DB 6 ISCNNSGKDGNTSANSANSADESVKGNLTETNKKITDSNAVLAVKEVEALLSIDETIAKA 65
| : |||||

QY 61 IGKRIHQNGLDPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAAKCSFTFN 120
|||||
DB 66 IGKRIHQNGLDPTENNHNSSLAGAYASTLIKQKIDGLKNEGLKEKIDAAKCSFTFN 125
|||||

QY 121 KLEKHTDIDGKEGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKA 170
| |||||
DB 126 KLEKHTDIDGKEGVTDADAKKAILKANGTKTKGAELGKLFESVEVLSKA 175
| |||||

RESULT 13
Q9R7A9 PRELIMINARY: PRT: 163 AA.

AC Q9R7A9 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
RT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT U.S.A.'";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91801; AAB81898.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1 163
FT SEQUENCE 163 AA; 17210 MW; 7550141651BD01FF CRC64;

Query Match 43.1%; Score 802; DB 2; Length 163;
 Best Local Similarity 98.2%; Pred. No. 4.8e-30;
 Matches 160; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVYKGNLTKITDSNAVLAVKEVALLSIDEIAAKAI 61
 DB 1 SCNNSGKDGNTSANSADSVYKGNLTKITDSNAVLAVKEVALLSIDEIAAKAI 60

QY 62 GKRIHONNGLDTEENNHSLLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFTTK 121
 DB 61 GKRIHONNGLDTEENNHSLLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFTTK 120

QY 122 LKEKHTDLGKEGYTDADAKKAILKANGTKTGAEELGKLFESV 164
 DB 121 LKEKHTDLGKEGYTDADAKKAILKANGTKTGAEELGKLFESV 163

LT 14
 ID 09AGBI PRELIMINARY; PRT; 165 AA.
 AC 09AGBI;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OSCP-MODIFIED.
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB19;
 RA Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
 RA Lawson C.L., Swaminathan S.;
 RT "Crystal structure of outer surface protein C (OspC) from the Lyme
 RT disease spirochete, Borrelia burgdorferi.";
 RL EMBL J. 0:0-0(2001).
 DR EMBL: AF37548: AAK21289.1; M -> I.
 FT VARIANT 61 61
 SQ SEQUENCE 165 AA; 17833 MW; 56E01536D22F61BF CRC64;

Query Match 42.3%; Score 786; DB 2; Length 165;
 Best Local Similarity 98.2%; Pred. No. 2.6e-29;
 Matches 161; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 215 VGPMLTEISKITTSNAVLAVKEVETLTSIDELAKAIGKRIKNDVSLDNEADHNSL 274
 DB 1 MGPMLTEISKITTSNAVLAVKEVETLTSIDELAKAIGKRIKNDVSLDNEADHNSL 60

QY 275 ISGAVLISMLTKRISAKIDSGELKAEIEKAKKCESEFTAKIKGHTDLGEGYTDNAK 334
 DB 61 MSGAVLISMLTKRISAKIDSGELKAEIEKAKKCESEFTAKIKGHTDLGEGYTDNAK 120

QY 335 KAILTNDKTKGADELKLFESVKNLSKAAKEMLTNSVKELTS 378
 DB 121 KAILTNDKTKGADELKLFESVKNLSKAAKEMLTNSVKELTS 164

RESULT 15
 ID 044720 PRELIMINARY; PRT; 211 AA.
 AC 044720;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OSCP.
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=25015;
 RX MEDLINE=95154673; PubMed=7851744;
 RA Stevenson B., Barthold S.W.;
 RT "Expression and sequence of outer surface protein C among North
 RT American isolates of Borrelia burgdorferi.";
 RL FEMS Microbiol. Lett. 124:367-372(1994).
 DR EMBL: U04282: AAC5540.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 211 AA; 22603 MW; 63984BA6D8743ED5 CRC64;

Query Match 41.7%; Score 774.5; DB 2; Length 211;
 Best Local Similarity 81.0%; Pred. No. 1.1e-28;
 Matches 158; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 1 MACNNSGKDGNTSANSADSVYKGNLTKITDSNAVLAVKEVALLSIDEIAAK 59
 DB 17 ISCNNSGKDGNTSANSADSVYKGNLTKITDSNAVLAVKEVALLSIDEIAAK 76

QY 60 AIGKKIHOONNGLDTEENNHSLLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFT 119
 DB 77 AIGKKIHOONNGLDTEENNHSLLAGAVASTLIKOKLDGKNEGLKEKIDAAKCSSEFT 136

QY 120 NKLKHTDLGKEGYTDADAKKAILKANGTKTGAEELGKLFESVYLSKAAKEMLANSV 179
 DB 137 NKLKSHTELGKQDDADDADAKKAILRTHTKGAEEIDKLFKAVENLSKAAKEMLSNV 196

QY 180 KELTSPVVAESPCKP 194
 DB 197 KELTSPVVAESPCKP 211

Search completed: March 18, 2002, 10:10:53
 Job time: 976 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:30 ; Search time 118.14 Seconds

(without alignments)
236.377 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850

Sequence: 1 MACNNSGKDCGTSANSADES.....AVENLAKAKEMIANSVKEL 377

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.1101.*

- 1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1850	100.0	377	22 AAB62713
2	1825	98.6	400	22 AAB62739
3	1796.5	97.1	373	22 AAB62711
4	1793.5	96.9	397	22 AAB62729
5	1614.5	87.3	378	22 AAB62712
6	1589.5	85.9	401	22 AAB62738
7	1561	84.4	374	22 AAB62710
8	1558	84.2	398	22 AAB62728
9	1519.5	82.1	369	22 AAB62716
10	1516.5	82.0	393	22 AAB62731
11	1480.5	80.0	367	22 AAB62719

12	1477.5	79.9	391	22 AAB62735	Borrelia sp chimera
13	1467	79.3	384	22 AAB62726	Borrelia sp chimera
14	1465	79.2	560	22 AAB62724	Borrelia sp chimera
15	1464	79.1	408	22 AAB62737	Borrelia sp chimera
16	1284	69.4	370	22 AAB62715	Borrelia sp chimera
17	1281	69.2	394	22 AAB62730	Borrelia sp chimera
18	1259.5	68.1	369	22 AAB62714	Borrelia sp chimera
19	1256.5	67.9	393	22 AAB62732	Borrelia sp chimera
20	1245	67.3	368	22 AAB62718	Borrelia sp chimera
21	1242	67.1	392	22 AAB62734	Borrelia sp chimera
22	1225.5	66.2	368	22 AAB62717	Borrelia sp chimera
23	1217.5	65.8	391	22 AAB62736	Borrelia sp chimera
24	1183.5	64.0	378	22 AAB62725	Borrelia sp chimera
25	1173.5	63.5	401	22 AAB62733	Borrelia sp chimera
26	1164	62.9	410	22 AAB62740	Borrelia sp chimera
27	1163	62.9	386	22 AAB62727	Borrelia sp chimera
28	972.5	52.6	466	16 AAR75740	B31 outer surface
29	954	51.6	209	22 AAB62720	B burgdorferi ospC
30	945	51.1	210	16 AAM11935	Outer surface prot
31	943	51.0	192	15 AAR60886	Borrelia IP2 OSpC
32	939	50.8	192	22 AAB62721	B burgdorferi ospC
33	929	50.2	587	16 AAR75746	B31 Osp-A/antigen
34	928	50.2	210	16 AAR75727	B. burgdorferi str
35	915	49.5	466	16 AAR75739	B31 outer surface
36	908	49.1	192	22 AAB62703	B burgdorferi ospC
37	876	47.4	193	22 AAB62709	B burgdorferi ospC
38	864	46.7	176	15 AAR62772	Borrelia B31 antlg
39	707	38.2	192	15 AAR60889	Borrelia 297 antlg
40	694.5	37.5	177	15 AAR62773	Borrelia 25015 ant
41	694.5	37.5	177	15 AAR60887	Borrelia 25015 osp
42	692.5	37.4	211	18 AAM41822	B burgdorferi sen
43	692.5	37.3	211	20 AAM93473	B. burgdorferi Osp
44	690.5	37.3	193	15 AAR60888	Borrelia Z57 antlg
45	685	37.0	212	16 AAM11934	Outer surface prot

ALIGNMENTS

RESULT 1	
AA62713	
ID	AA62713 standard; Protein; 377 AA.
XX	
AC	AA62713;
XX	
DT	03-APR-2001 (first entry)
XX	
DE	Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX	
KW	Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX	
OS	Chimeric - Borrelia sp.
XX	
PN	Chimeric - Borrelia sp.
XX	
PD	WO200078966-A1.
XX	
PF	28-DEC-2000.
XX	
PR	19-JUN-2000; 2000MO-US16915.
XX	
XX	18-JUN-1999; 99US-0140042.
XX	
PA	(UYNT) UNIV NEW YORK STATE RES FOUND.
XX	
PI	(BROO-) BROOK BIOTECHNOLOGIES INC.
XX	
DR	Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX	
XX	WPI: 2001-050113/06.
XX	N-PSDB; AAR29017.
PT	Compositions of ospC polypeptides from strains of Borrelia which cause
PT	Lyme disease are used to immunize animals and detect immune responses
PT	to Lyme disease -

XX Claim 43; Page 86-87; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

XX Sequence 377 AA;

Query Match 100.0%; Score 1850; DB 22; Length 377;
 Best Local Similarity 100.0%; Pred. No. 1,6e-124;
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNLTETINKRTDSNNAVLAWEVALLSSIDEIAAKA 60
 1 macnsgkdgntsansadesvkgpnlteinkrtidsnavllaaveallssideiaaka 60
 QY 61 IGGKTHONNGLDTENNHNGSLLAGAVAI STLKOKLDGLKNGLEKIDAKKCEFTFN 120
 61 igkthongldtennhngsllagayaistlikqldglneglekeidaakksetftn 120
 Db 61 IGGKTHONNGLDTENNHNGSLLAGAVAI STLKOKLDGLKNGLEKIDAKKCEFTFN 120
 QY 121 KLEKHTDGLKEGYTDADAKFAILKANGTKTGAELGKLFESVEVLSKAKEMLANSVK 180
 121 klekhtdlgkegytdadakeaailkangtkkgaeeigklfesevylskaakemlansvk 180
 Db 121 KLEKHTDGLKEGYTDADAKFAILKANGTKTGAELGKLFESVEVLSKAKEMLANSVK 180
 QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKRTTESNAVLAWEKEI 240
 181 eltspvvaesppkpsmvnnsqgdntsansadesvkgpnlteiskrttesnavlavwekei 240
 Db 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKRTTESNAVLAWEKEI 240
 QY 241 ETLASIDELATKATGKTIQONGGLAVEAGHNGTLAAGAVTISKLTOKLDGLKNSKELK 300
 241 etllasidelatkaigkiiqongglaveaghnngtllagaytiskltitqldglnskelk 300
 Db 241 ETLASIDELATKATGKTIQONGGLAVEAGHNGTLAAGAVTISKLTOKLDGLKNSKELK 300
 QY 301 EKIENAKKCEDEFTKLEGEHAQIGIENVTDENAKKATLITDAKKGAAELKLFKAYE 360
 301 ekienakkcedftkkllegehaqigienvtndenakkaillitdaakkgaaeleklfkaye 360
 Db 301 EKIENAKKCEDEFTKLEGEHAQIGIENVTDENAKKATLITDAKKGAAELKLFKAYE 360
 QY 361 NLAKAKEMLANSVKEL 377
 361 nlakaakemlansvkel 377
 Db 361 nlakaakemlansvkel 377

RESULT 2
 AAB62739

AC AAB62739 standard; Protein; 400 AA.

XX AAB62739;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYNY) UNIV NEW YORK STATE RES FOUND.

PI (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29043.

XX Compositions of ospc polypeptides from strains of Borrelia which cause

PT Lyme disease are used to immunize animals and detect immune responses

PT to Lyme disease -

XX Claim 43; Page 153; 160pp; English.

PS The present invention provides compositions comprising ospc proteins and

CC chimeric ospc proteins from members of the Borrelia genus. These may be

CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as

CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 400 AA;

Query Match 98.6%; Score 1825; DB 22; Length 400;
 Best Local Similarity 98.9%; Pred. No. 1,1e-122;
 Matches 372; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTDSNNAVLAWEVALLSSIDEIAAKAI 61
 2 acnnsqgdntsansadesvkgpnlteinkrtidsnavllaaveallssideiaakai 61
 Db 25 scnsgkdgntsansadesvkgpnlteiskrtidsnavllaaveallssideiaakai 84
 QY 62 GKTIHONNGLDTENNHNGSLLAGAVAI STLKOKLDGLKNGLEKIDAKKCEFTFNK 121
 62 gkthongldtennhngsllagayaistlikqldglneglekeidaakksetftnk 121
 Db 85 gkthongldteyhnngsllagayaistlikqldglneglekeidaakksetftnk 144
 QY 122 LKEKHTDGLKEGYTDADAKFAILKANGTKTGAELGKLFESVEVLSKAKEMLANSVKE 181
 122 lkekhtdlgkegytdadakeaailkangtkkgaeeigklfesevylskaakemlansvke 181
 Db 145 lkekhtdlgkegytdadakeaailknglctkgaeeigklfesevylskaakemlansvke 204
 QY 182 LTSPPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKRTTESNAVLAWEKEI 241
 182 ltsppvvaesppkpsmvnnsqgdntsansadesvkgpnlteiskrttesnavlavwekei 241
 Db 205 ltsppvvaesppkpsmvnnsqgdntsansadesvkgpnlteiskrttesnavlavwekei 264
 QY 242 TLLASIDELATKATGKTIQONGGLAVEAGHNGTLAAGAVTISKLTOKLDGLKNSKELKE 301
 242 tllasidelatkaigkiiqongglaveaghnngtllagaytiskltitqldglnskelke 301
 Db 265 tllasidelatkaigkiiqongglaveaghnngtllagaytiskltitqldglnskelke 324
 QY 302 KIENAKKCEDEFTKLEGEHAQIGIENVTDENAKKATLITDAKKGAAELKLFKAYEN 361
 302 kienakkcedftkkllegehaqigienvtndenakkaillitdaakkgaaeleklfkayen 361
 Db 325 kienakkcedftkkllegehaqigienvtndenakkaillitdaakkgaaeleklfkayen 384
 QY 362 LAKAKEMLANSVKEL 377
 362 lakaakemlansvkel 377
 Db 385 lakaakemlansvkel 400

RESULT 3

AC AAB62711 standard; Protein; 373 AA.

XX AAB62711;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 26.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

XX (UINY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX
PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI: 2001-050113/06.
DR N-PSDB; AAF29015.
DR
XX
XX
XX Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX
PS Claim 43: Page 81: 160pp: English.
XX
XX
XX The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

Sequence 373 AA;

Query Match	97.18;	Score 1796.5;	DB 22;	Length 373;
Best Local Similarity	98.18;	Pred. No. 1e-120;		
Matches 369; Conservative	2;	Mismatches 2;	Indels 3;	Gaps 1;

0Y	2	ACNNSGKDNTSANSADSESVGPNLTPEINKITFDSNAYLLAVEVEALLSIDIEIAAKI	61
Db	1	acmnsygdntsansadesvkgpnltetstskkltidsnavllaveveallsideiaekal	60
0Y	62	GKKITHONGLDPTENNHNHSGSLAGAYASTLTIKOKLDLKGKLEKEDIIDAAKCSFTPNK	121
Db	61	gkltihngmldpteynhngsllagayastlilkpjlldgllkneglkexidaakcsftfntk	120
0Y	122	LKEHNTDLGKSGVYDDADAKKAILKANSTKTRKGAELEKLEESVEVLSAKKEMLANSYKE	181
Db	121	lkexhtldlkgvycdadakealltkngtlckkgaeelgkltiesvevlskaakemlansyke	180
0Y	182	LTPSVVAESPKKPSPMVNNSGKDNTSANSADSESVKGNLTREISKITTESNAYVLAYKEIE	241
Db	181	ltspvsves---pamvnsygdntsansadesvkgpnltetstskltiesnavylaveje	237
0Y	242	TLLASIDELATKAIGKRIQONGTSLAVEAGHNGTLLAAYVTSKLTITOKLDGLKNSKELKE	301
Db	238	tlslsldelatekælgkrltqnggslaveaghnngtllagayltstskltitqklgllknselke	297
0Y	302	KIENNAKCSDEFTKKLEGEHAQCIENVNTDENAKKAILITDAADKDKGALEKLEKFAVEN	361
0Y	298	kienakcsedftckllegehaqglienvntdenakkailitdaadkdkgaalelekifaven	357
0Y	362	LAKAKEMLANSYKEL 377	
Db	358	lakakemlansykel 373	

RESULT 4

ID AAB62729 standard; Protein; 397 AA.

AC AAB62729;

DT 03-APR-2001 (first entry)

Borrelia sp chimeric ospC protein SEQ ID NO: 62.

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

Accession	Strain	Host	Location	Year
05	Chimeric	Borrelia sp.		
05	Chimeric	Borrelia sp.		

PN WO2000078966-A1.

XX 28-DEC-2000
PD

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UUNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR

XX
XX
concerns of the Council from establishment of Borrowia which caused

PT to Lyme disease

PS Claim 43; Page 125-126; 160pp; English

CC The present invention provides compositions comprising ospc proteins and

CC vaccines against *Borrelia* infection, which is spread by ticks and leads
CC to Lyme disease.

SQ **Sequence** **397** **AA;**

Query Match	96.9%	Score 1793.5	DB 22	Length 397
Best Local Similarly	97.9%	Pred. No. 1.9e-120		
Matches 368	Conservative	3	Mismatches 2	Indels 3
				Gaps 1

QY 2 ACNNSGKDNTSANSADSEVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61

Db 25 scnsqkdntsansadesvkgpnlteiskittdsnavlavkevealssidelaaka1 84

62 GKKIHQNNGLDTENNHNGLAGAYAISTLLIKQKLDGLKNEGLKEKIDAKKCEFTNK 121

Db 85 gkklhngngldteynhngslJagaya1stllKqklDgJkneglkeklDaakkcsetltK 144

122 LKERHIDLGKEGVJDADAKKAILKANGIKIKGAELGKLFESVELSKAAKEMILANSVKE 181

D5 145 LKekhtdlgkegvCdaadakea1LktngrKrkgaee1gkLtesVeV1SKraakem1ansvKe 204

182 LISPVVAESPKKPSMVNNSGDGNISANSADESVAGPNLEISKIIESNAVLAVKEIE 241

DD 203 LSPVdaes - - - palivlinisgkugilcsaaisdaesvkypilceiskllesinavlavkele 201

[illegible]

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

100

QY	362	LAKAKEMLANSVKEL	377
Db	382	lakaakemlansvkel	397

RESULT	5
RESULT	5

ID	standard; Protein; 378 AA.
AAB62712	

AC AAB62712;

DT 03-APR-2001 (first entry)

DE Borrella sp chimeric ospc protein SEQ ID NO: 28.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX W0200078966-A1.
 XX 28-DEC-2000.
 XX 19-JUN-2000; 2000MO-US16915.
 XX 18-JUN-1999; 99US-0140042.
 XX (UANY) UNIV NEW YORK STATE RES FOUND.
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 PI WPI: 2001-050113/06.
 XX N-PSDB; AAF29016.
 XX Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX Claim 43; Page 83-84; 160pp; English.
 PS The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX Sequence 378 AA;
 SQ

Query Match 87.3%; Score 1614.5; DB 22; Length 378;
 Best Local Similarity 87.8%; Pred. No. 1.1e-107;
 Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGPNTLEINRKTTSNAVLAVKEVEALLSSIDEIAAKA 60
 DB 1 macnsgkdgtansadesvkgpnlteinkktidseavllavkeveallssideiaaka 60
 QY 61 IGRKHONNGLDTENNHNSILAGAVAISTLIKOKIDGKNEGKLEKIDAAKCEFTFN 120
 DB 61 igkkihqngldtenhngsllagayaistlikqkldgkneglekidaakcseftfn 120
 QY 121 KLEKHTDGLKEGVTDADAKKAILKANGRTKGADELGLFESVEVLSKAKEMLANSVK 180
 DB 121 klekhtdglkegvtdadakeailkangrltkgaeelglfesevevlskakemlansvk 180
 QY 181 ELTSPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEI 240
 DB 181 eltspvvaespckpsmvnnskgdntsansadesvkgpnlteiskttessnavlavkei 240
 QY 241 ETLASIDELATKAIGKTIQONGSLAVEAGHNGTLGAVTISKLITOKIDGLKNSKTK 300
 DB 241 eltllsidedela-kaigkktikndvsldneadngsllsagaylislitkksaidsgekl 300
 QY 301 EKIEAKKCEDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGALEKLEKRAVE 360
 DB 300 aeieakkcseeflaklgehtdylgegvtdadnakkailitndktkgadeleklelesvk 359
 QY 361 NLAAKAKEMLANSVKEL 377
 DB 360 nlksaakemltnsvkel 376

RESULT 6
 ID AAB62738
 XX AAB62738 standard; Protein; 401 AA.

AC AAB62738;
 XX 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 80.
 XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 KW Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX W0200078966-A1.
 XX 28-DEC-2000.
 XX 19-JUN-2000; 2000MO-US16915.
 XX 18-JUN-1999; 99US-0140042.
 XX (UANY) UNIV NEW YORK STATE RES FOUND.
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 PI WPI: 2001-050113/06.
 XX N-PSDB; AAF29042.
 XX Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX Claim 43; Page 150-151; 160pp; English.
 PS The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX Sequence 401 AA;
 SQ

Query Match 85.9%; Score 1589.5; DB 22; Length 401;
 Best Local Similarity 86.7%; Pred. No. 7.2e-106;
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEINRKTTSNAVLAVKEVEALLSSIDEIAAKAI 61
 DB 2 acnnskgdgtansadesvkgpnlteiskttidseavllavkeveallssideiaakai 61
 QY 62 IGRKHONNGLDTENNHNSILAGAVAISTLIKOKIDGKNEGKLEKIDAAKCEFTFN 121
 DB 62 igkkihqngldtenhngsllagayaistlikqkldgkneglekidaakcseftfn 121
 QY 122 KLEKHTDGLKEGVTDADAKKAILKANGRTKGADELGLFESVEVLSKAKEMLANSVK 181
 DB 122 klekhtdglkegvtdadakeailkangrltkgaeelglfesevevlskakemlansvk 181
 QY 181 ELTSPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPNTLEISKTTESNAVLAVKEI 241
 DB 181 eltspvvaespckpsmvnnskgdntsansadesvkgpnlteiskttessnavlavkei 241
 QY 241 ETLASIDELATKAIGKTIQONGSLAVEAGHNGTLGAVTISKLITOKIDGLKNSKTK 301
 DB 241 eltllsidedela-kaigkktikndvsldneadngsllsagaylislitkksaidsgekl 301
 QY 301 EKIEAKKCEDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGALEKLEKRAVE 361
 DB 302 aeieakkcseeflaklgehtdylgegvtdadnakkailitndktkgadeleklelesvk 363
 QY 362 NLAAKAKEMLANSVKEL 377
 DB 384 lskaakemltnsvkel 399

Query Match	84.2%	Score 1558;	DB 22;	Length 398;
Best Local Similarity	85.6%	Pred. No. 1.3e-103;		
Matches 322;	Conservative 23;	Mismatches 27;	Indels 4;	Gaps 2;
2	AANNNGKOSNTSANSADSVKGNPNTEIKKTTDSNAVLAVKEVFAALLSSIDEINAKAI	61		
25	scnnsgkdgntsansadsdvkgnptltskltldsnavlavegaallssideiaakai	84		
62	GKRIQONNGLDENNNHNSGLAGAVAIISYLIROKLDGLNKGKTEIDAAKCSFTYTK	121		
85	gkllqngngldceyhngslagayaisltllqkldglnknegleidaakkcsftfk	144		
122	LKEKHTDLCGECVTDADAKKAILKANGTYTKGAEBELGKLFESVEVLSKAKEMLANSVKE	181		
145	lkehthldlgkegvtdadakeaillktngctkgaeeigklifesvevlskaakemlansvke	204		
182	LTSPPVABSPKPSWVNNNGKDGNTSANSADSVKGNPNTEISKTTESNAVVLAVKEIE	241		
205	ltsppvaee--panwnnsgkdgntsansadsdvkgnptltskltldsnavlavega	261		

```

QY 242 TLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTISKLITQKLDGLKNSSEKLE 301
DB 262 TLITSIDEla-kaigkiknovslidneadngslisgayllsnlltkksaidsgeleka 320
QY 302 KIENAKKSEDFPKKLEGEHAOLGIENVTDENAKKALLITDAKDKGALEKLEKFRAYEN 361
DB 321 elefakkseeftaklgehtldlgkevtdnakkalkknndkttgadeloklifesvkn 380
QY 362 LAKAKKEMLANSVKEL 377
DB 381 lskakemltnsvkel 396

RESULT 9
AAB62716
ID AAB62716 standard; Protein; 369 AA.
AAB62716:
DT 03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 36.
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX WO200078966-A1.
XX 28-DEC-2000.
XX 19-JUN-2000; 2000WO-US16915.
XX 18-JUN-1999; 99US-0140042.
XX (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI; 2001-050113/06.
XX DR N-PSDB; AAF29020.
XX PA Compositions of Ospc polypeptides from strains of Borrelia which cause
XX PT Lyme disease are used to immunize animals and detect immune responses
XX PI to Lyme disease -
XX PS Claim 43; Page 94-95; 160pp; English.
XX CC The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX SQ Sequence 369 AA:

Query Match 82.1%; Score 1519.5; DB 22: Length 369;
Best Local Similarity 86.2%; Pred. No. 6.5e-101;
Matches 326; Conservative 10; Mismatches 31; Indels 11; Gaps 4;
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTIDSNVAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 acnnsykgdntsansadesvkgpnlteiskrtidsnavllavkeveallssidel-akai 59
QY 62 GKTIHONGGIDTENNHNHNSLAGAVALSTLKQKLDGLK-NEGKLEKIDAKKCKSEFTTN 120
DB 60 gkklndgslidneanrnesllagayllstltqklsklnsgelikekfaakckseest 119
QY 121 KLKKEHTDGLKEGVTDADAKPAILKANGT-KTGAEELGKLFESVEVLAKAKKEMLANSV 179

```

```

DB 120 klkdnhaqligvtdenakkalkkanaagkdkgyeeleklsgslslejskaakemlansv 179
QY 180 KELTSPVAVESPKRPMNVNNSGKDGNTSANSADSVKGNLTETINKRTIDSNVAVLAVKE 239
DB 180 keltspvavg-----nmsgkdntsansadesvkgpnlteiskrtidsnavllavke 231
QY 240 IETLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTISKLITQKLDGLKNSSEK 299
DB 232 ietllasidelatkaigkligqnglaveagngtllagaytlstklitqkldglknsel 291
QY 300 KEKIEAKKSEDFPKKLEGEHAOLGIENVTDENAKKALLITDAKDKGALEKLEKFRAY 359
DB 292 kekienakkseefcklegehaqligvtdenakkallitdaakdkgyaealeklfrav 351
QY 360 ENLAKAKKEMLANSVKEL 377
DB 352 enlakaakemlansvkel 369

RESULT 10
AAB62731
ID AAB62731 standard; Protein; 393 AA.
AC AAB62731;
XX 03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 66.
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX WO200078966-A1.
XX 28-DEC-2000.
XX 19-JUN-2000; 2000WO-US16915.
XX 18-JUN-1999; 99US-0140042.
XX (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI; 2001-050113/06.
XX DR N-PSDB; AAF29035.
XX PA Compositions of Ospc polypeptides from strains of Borrelia which cause
XX PT Lyme disease are used to immunize animals and detect immune responses
XX PI to Lyme disease -
XX PS Claim 43; Page 131; 160pp; English.
XX CC The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX SQ Sequence 393 AA:

Query Match 82.0%; Score 1516.5; DB 22: Length 393;
Best Local Similarity 86.0%; Pred. No. 1.2e-100;
Matches 325; Conservative 11; Mismatches 31; Indels 11; Gaps 4;
QY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTIDSNVAVLAVKEVEALLSSIDEIAAKAI 61
DB 25 acnnsykgdntsansadesvkgpnlteiskrtidsnavllavkeveallssidel-akai 83

```


QY 62 GKKIHONGDTENNHNGSLAGAVATSTLIKOKLOGLK-NEGKKEKIDAAKCESETFN 120
 DB 84 gkikndgsldneenresllagaytsclicqlskingseglkexiaakccseefst 143
 QY 121 KLEKHTDLCEGYVDAAKEALIKANGT-KTGAELGLFESVEVLSKAEMLANSV 179
 DB 144 klndhagqiglytdeenakalkkanaagkdkyveeleklsglslesksaakemlansv 203
 QY 180 KELTSPVAESPCKPSMVNNSGKDGMTSANSADSVKGPMLTEISKITTESNAVLAKE 239
 DB 204 keltspvvhg-----nsgkdgnatsansadesvkgpmltelskkitcsnavlavke 255
 QY 240 IETLLASIDELATKATGKTIQONGGLAVEGHNGTLLAGAVTTSKITOKLIDGKNSKEXL 299
 DB 256 Ietllasidelatkaigkxiqngglaveagngcllagaytsclicqlidgiknsek1 315
 QY 300 KEKIEAKKCEDEFTKLEGEHAQLGIENVYDENAKKAILITDAKDKGAELKFLKAV 359
 DB 316 kekienakccsedftkklgeghagqiglytdeenakalkltdaakdkgaaleklfkav 375
 DB 360 ENLAKAKEMLANSVKEL 377
 DB 376 enlakaakemlansvkel 393

RESULT 11
 AAB62719
 ID AAB62719 standard; Protein: 367 AA.
 AC AAB62719;
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 42.
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 PD 28-DEC-2000.
 PF 19-JUN-2000; 2000WO-US16915.
 PR 18-JUN-1999; 99US-0140042.
 PA (UYNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI: 2001-050113/06.
 DR N-PSDB: AAF29023.
 PT Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 PS Claim 43; Page 102-103; 160pp; English.
 XX- The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX Sequence 367 AA;
 SQ

Query Match 80.0%; Score 1480.5; DB 22; Length 367;

Best Local Similarity 82.4%; Pred. No. 3,9e-98;
 Matches 310; Conservative 23; Mismatches 34; Indels 9; Gaps 2;
 QY 2 ACNNSGKDGMTSANSADSVKGPMLTEINKITDSNAVLAKEVEALLSIDELAKAI 61
 DB 1 acnnskdgnatsansadesvkgpmltelskkitcsnavlavkeveltsidelakai 60
 QY 62 GKKIHONGDTENNHNGSLAGAVATSTLIKOKLOGKNGLEKIDAAKCESETFNK 121
 DB 61 gkikndgsldneenresllagaytsclicqlskingseglkexiaakccseefst 119
 QY 122 KLEKHTDLCEGYVDAAKEALIKANGT-KTGAELGLFESVEVLSKAEMLANSVKE 181
 DB 120 kleshavlgidnldcnaagralikkhankdkgaaleklfkavenlkaagdtcknavke 179
 QY 182 LTSPVAESPCKPSMVNNSGKDGMTSANSADSVKGPMLTEISKITTESNAVLAKEIE 241
 DB 180 ltspvvhg-----nsgkdgnatsansadesvkgpmltelskkitcsnavlavkeie 231
 QY 242 TLLASIDELATKATGKTIQONGGLAVEGHNGTLLAGAVTTSKITOKLIDGKNSKEXLKE 301
 DB 232 tllasidelatkaigkxiqngglaveagngcllagaytsclicqlidgiknsek1ke 291
 QY 302 KIENAKKCEDEFTKLEGEHAQLGIENVYDENAKKAILITDAKDKGAELKFLKAVEN 361
 DB 292 kienakccsedftkklgeghagqiglytdeenakalkltdaakdkgaaleklfkaven 351
 QY 362 LAKAKEMLANSVKEL 377
 DB 352 lakaakemlansvkel 367

RESULT 12
 AAB62735
 ID AAB62735 standard; Protein: 391 AA.
 AC AAB62735;
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 74.
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 PD 28-DEC-2000.
 PF 19-JUN-2000; 2000WO-US16915.
 PR 18-JUN-1999; 99US-0140042.
 PA (UYNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI: 2001-050113/06.
 DR N-PSDB: AAF29039.
 PT Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 PS Claim 43; Page 142; 160pp; English.
 XX- The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.
XX
SQ Sequence 391 AA:

Query Match 79.9%; Score 1477.5; DB 22; Length 391;
Best Local Similarity 82.2%; Pred. No. 7e-98;
Matches 309; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

```

OY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTDSNAVILAVKEVEALLSSIDEIAAKAI 61
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 25 scmsgkdgntansadesvkgpnlteiskitdsnavilavkevelssideiaakai 84
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 62 GKIIHONNGLDTEENNHNGLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCEFTFNK 121
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 85 gkkl-gnglgeangekntslsgayaisdliaeklnvklneelkexidtaqgstetfnk 143
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 122 LKEKHTDGLKEGVTADADKAEATLKANGTKYTGAEELGKLFESVEVLSKAAKEMLANSVKE 181
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 144 lksnavlglndltddagralikkkhankdgyaaeleklfkavenlskaaqdtlknave 203
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 182 LTSPVVAESPKKPSKVNNSGKDGNTSANSADSVKGNLTETISKITEESNAVILAVKEIE 241
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 204 ltsplvbg-----nmsgkdgntansadesvkgpnlteiskitdsnavilavkeie 255
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 242 TLASIDELATKAI GKTIQONGGLAVAGHNGTLLAGAVTISKLITQKLDGLKNSKELKE 301
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 256 tllasidelakai gkklqngglavaghnngtllagaytiskliltqklidglnseklike 315
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 302 KIENAKKCEDEFTKLEGEHAQLGIENVTDENAKKATLITDAKDKGAELKELFKAVEN 361
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 316 kienakkcsedfckklegehaqgienvtdenakkallitdaakdkgyaaeleklfkaven 375
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 362 LAKAAKEMLANSVKEL 377
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 376 lakaakemlansvkel 391
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 13

AAB62726
ID AAB62726 standard; Protein; 384 AA.

XX AAB62726;
XX
DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 56.

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwlyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

XX N-PSDB; AAF29030.

XX Compositions of ospC polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 117-118; 160pp; English.

CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. atzeli or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

XX
SQ Sequence 384 AA:

Query Match 79.3%; Score 1467; DB 22; Length 384;
Best Local Similarity 81.4%; Pred. No. 3.9e-97;
Matches 306; Conservative 30; Mismatches 36; Indels 4; Gaps 3;

```

OY 2 ACNNSGKDGNTSANSADSVKGNLTETINKRTDSNAVILAVKEVEALLSSIDEIAAKAI 61
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 62 GKIIHONNGLDTEENNHNGLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCEFTFNK 121
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 gkklngngldteyhnngslagayaistllkqklidglnknglkekidaakcetfnk 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 122 LKEKHTDGLKEGVTADADKAEATLKANGTKYTGAEELGKLFESVEVLSKAAKEMLANSVKE 181
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 121 lkehtdglkegvtadadkaealktngctkgyaaelgklfsvavlksaakemlansvke 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 182 LTSPVVAESPKKPSKVNNSGKDGNTSANSADSVKGNLTETISKITEESNAVILAVKEIE 241
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 181 ltspyvaesppkkpfngnsg--gdsasntpdesakgplvtvtskklidsnavilavve 238
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 242 TLASIDELATKAI GKTIQONGGLAVAGHNGTLLAGAVTISKLITQKLDGLKNSKELKE 301
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 239 allsidel-skaigkklkngdltldneanresllagayelskliltqklv1-nseelk 296
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 302 KIENAKKCEDEFTKLEGEHAQLGIENVTDENAKKATLITDAKDKGAELKELFKAVEN 361
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 297 kikeakdcsgkftcklkksheajlgisvqddnakalltkhtgkdkagakeleelksles 356
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 362 LAKAAKEMLANSVKEL 377
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 357 lskaagaallnsvkel 372
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 14

AAB62724
ID AAB62724 standard; Protein; 560 AA.

XX AAB62724;
XX
DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 52.

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwlyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

DR N-PSDB: AAF29028.

XX Compositions of OspC polypeptides from strains of *Borrelia* which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease.

PS Claim 43: Page 112-113; 160pp; English.

XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the *Borrelia* genus. These may be
CC *Borrelia burgdorferi*, *B. afzelii* or *B. garinii*. These can be used as
CC vaccines against *Borrelia* infection, which is spread by ticks and leads
CC to Lyme disease.

XX Sequence 560 AA:

Query Match 79.2%; Score 1465; DB 22; Length 560;
Best Local Similarity 81.0%; Pred. No. 8, 8e-97;
Matches 306; Conservative 23; Mismatches 45; Indels 4; Gaps 2;

1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAAKA 60
|||||
1 macnsgkdgnstansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 60

QY 61 IGGKIHQNNGLDTENNHNHNSLAGAVAIITLIKOKLDGLKNEGKIDAKKCSSEFTFN 120
|||||

DB 61 Iggkihqnnngldteyhnngslagayaisltikqldgiknegikeidaakkcsseftfn 120

QY 121 KLEKHTDGLGEGVTADAKKAIKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
|||||

DB 121 klekhdtlgyegvtadakeaalkngtktkgaeelgklfesvevlskaakemlansvk 180

QY 181 ELTSPVVAESPKKPSMVNNSGKDGNT-SANSADSVKGNLTETINKKITESNAVLAVKE 239
|||||

DB 181 eltspvvae---pangnsgksgdsastnpadesakgnlteiskkitdsnavlavke 237

QY 240 IETLSTIDELATKATGKTIQONGSLAVAGHNGTLLAAGATYISKITOKLDGKNSKLE 299
|||||

DB 238 velivstidelakkaigkldnnglaalnqngslagayaisltleikslklnleel 297

QY 300 KKEIEAKKCEDEFTKLEGEHAQGIENVTDENAKKAILITPAKDKGALEKLFKAV 359
|||||

DB 298 kkeiakkkcseefnklisgshadlygqdatdhnakaalktchatcdkakeftkdlfesy 357

QY 360 ENLAKAKEMLANSVKEL 377
|||||

DB 358 egllkaagvaltnsvkel 375

UNT 15

ID AAB62737 standard; Protein; 408 AA.

AC AAB62737;

DT 03-APR-2001 (first entry)

DE *Borrelia* sp chimeric ospC protein SEQ ID NO: 78.

XX *Borrelia*; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - *Borrelia* sp.

XX Chimeric - *Borrelia* sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000MO-US16915.

PR 18-JUN-1999; 99US-0140042.

XX

PA (UNYK) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29041.

PT Compositions of OspC polypeptides from strains of *Borrelia* which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease.

PS Claim 43: Page 147-148; 160pp; English.

XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the *Borrelia* genus. These may be
CC *Borrelia burgdorferi*, *B. afzelii* or *B. garinii*. These can be used as
CC vaccines against *Borrelia* infection, which is spread by ticks and leads
CC to Lyme disease.

XX Sequence 408 AA:

Query Match 79.1%; Score 1464; DB 22; Length 408;
Best Local Similarity 81.1%; Pred. No. 6, 8e-97;
Matches 305; Conservative 31; Mismatches 36; Indels 4; Gaps 3;

QY 2 ACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAAKA 61
|||||

DB 25 cnsngkdgnstansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 84

QY 62 GKTIHNNGLDTENNHNHNSLAGAVAIITLIKOKLDGLKNEGKIDAKKCSSEFTFN 121
|||||

DB 85 gkkihqnnngldteyhnngslagayaisltikqldgiknegikeidaakkcsseftfn 144

QY 122 LKEKHTDGLGEGVTADAKKAIKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 181
|||||

DB 145 lkekhdtlgyegvtadakeaalkngtktkgaeelgklfesvevlskaakemlansvk 204

QY 182 LTSPPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETINKKITESNAVLAVKE 241
|||||

DB 205 ltspvvae---pangnsgksgdsastnpadesakgnlteiskkitdsnavlavke 262

QY 242 TLLASTIDELATKATGKTIQONGSLAVAGHNGTLLAAGATYISKITOKLDGKNSKLE 301
|||||

DB 263 allssidel-skaigkikndgldneannesiagayaisltkqlsvl-nseelk 320

QY 302 KIENAKKCEDEFTKLEGEHAQGIENVTDENAKKAILITPAKDKGALEKLFKAVEN 361
|||||

DB 321 kkeiakdcsgtktklksdhaeligsygdadnakkaalktchatcdkakeelkfxles 380

QY 362 LAKAKEMLANSVKEL 377
|||||

DB 381 lkaagaaltnsvkel 396

Search completed: March 18, 2002, 09:54:31
Job time: 329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:32 ; Search time 55.5 Seconds
(without alignments)
152.860 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSNADES.....AVENLAKAKEMIANSVKEL 377

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	972.5	52.6	466	4	US-08-235-836C-110 Sequence 110, App
2	945	51.1	210	1	US-08-158-353-3 Sequence 3, App1
3	938	50.7	210	4	US-08-209-603E-15 Sequence 15, App1
4	933	50.4	209	4	US-09-196-293-15 Sequence 15, App1
5	932	50.4	210	4	US-08-235-836C-30 Sequence 30, App1
6	929	50.2	588	4	US-08-235-836C-122 Sequence 122, App
7	915	49.5	466	4	US-08-158-353-2 Sequence 107, App
8	685	37.0	212	1	US-08-158-353-2 Sequence 2, App1
9	670	36.2	212	1	US-08-031-295-2 Sequence 2, App1
10	670	36.2	212	1	US-07-903-580-2 Sequence 2, App1
11	631	34.1	212	1	US-08-158-353-4 Sequence 4, App1
12	629	34.0	212	4	US-09-196-293-11 Sequence 11, App1
13	629	34.0	212	4	US-08-209-603E-11 Sequence 11, App1
14	626	33.8	212	4	US-08-235-836C-34 Sequence 34, App1
15	614.5	33.2	209	4	US-08-235-836C-32 Sequence 32, App1
16	592.5	32.0	207	4	US-08-235-836C-36 Sequence 36, App1
17	171	9.2	194	4	US-09-364-083-2 Sequence 2, App1
18	168	9.1	630	4	US-08-973-462-9 Sequence 9, App1
19	164.5	8.9	1164	4	US-08-923-992A-2 Sequence 2, App1
20	162.5	8.8	1164	4	US-08-923-992A-10 Sequence 10, App1
21	160	8.5	1786	4	US-08-923-992A-8 Sequence 8, App1
22	158	8.5	1098	4	US-08-923-992A-8 Sequence 8, App1
23	154.5	8.4	1128	4	US-08-923-992A-6 Sequence 6, App1
24	151.5	8.2	1104	4	US-08-923-992A-4 Sequence 4, App1
25	148	8.0	3111	2	US-08-460-309-4 Sequence 4, App1
26	148	8.0	3111	2	US-08-125-077-4 Sequence 4, App1
27	147.5	8.0	2482	1	US-08-328-254-6 Sequence 6, App1

28	146.5	7.9	688	3	US-09-141-047-8 Sequence 8, App1
29	146.5	7.9	3248	1	US-08-353-700-1 Sequence 1, App1
30	146.5	7.9	3248	5	PCT-US95-16216-1 Sequence 1, App1
31	145	7.8	1388	2	US-08-685-576-1 Sequence 1, App1
32	143.5	7.8	1388	2	US-08-685-576-4 Sequence 4, App1
33	143	7.7	414	5	PCT-US93-03077-3 Sequence 3, App1
34	143	7.7	1093	5	PCT-US93-03077-1 Sequence 1, App1
35	140.5	7.6	1196	1	US-08-144-121-4 Sequence 4, App1
36	140.5	7.6	1196	2	US-08-735-893-4 Sequence 4, App1
37	140	7.6	1147	3	US-08-470-260-5 Sequence 5, App1
38	140	7.6	1147	3	US-08-471-491-5 Sequence 5, App1
39	140	7.6	1147	4	US-08-466-662-5 Sequence 5, App1
40	140	7.6	3289	2	US-08-477-451-2 Sequence 2, App1
41	139.5	7.5	396	1	US-08-430-024-2 Sequence 2, App1
42	139.5	7.5	396	1	US-08-782-009-2 Sequence 2, App1
43	139.5	7.5	396	3	US-09-017-302-2 Sequence 2, App1
44	133.5	7.2	1354	3	US-08-685-871-2 Sequence 2, App1
45	132	7.1	2285	4	US-09-308-375-2 Sequence 2, App1

ALIGNMENTS

RESULT 1
US-08-235-836C-110
; Sequence 110, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF SEQUENCES: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-836C-110
Query Match 52.6%; Score 972.5; DB 4; Length 466;
Best Local Similarity 62.9%; Pred. No. 5.8e-66;
Matches 224; Conservative 23; Mismatches 74; Indels 35; Gaps 8;
QY 1 MACNNSGKDGNTSANSNADESVPKPNLTIKKITTDNAVLAVKVEALLSIDEIARA 60

```

Db 17 ISCNNSGKDGNTSANSADSVYKGNLTETISKRITDSNAVLLAVKEVALLSSIDEIAKA 76
      :|||||
Qy 61 ICKIHHNNGLDTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 120
      :|||||
Db 77 ICKIHHNNGLDTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 136
      :|||||
Qy 121 KLEKHTDGLKEGVTDADAKAEALILKANGTKTGAEELGKLFESVEVLSAAKEMLANSVK 180
      :|||||
Db 137 KLEKHTDGLKEGVTDADAKAEALILKANGTKTGAEELGKLFESVEVLSAAKEMLANSVK 196
      :|||||
Qy 181 ELTSPVVAESPKKPQ--VSSLD--EKNSVVDLPCEMKVLSKKNKDKGY----- 245
      :|||||
Db 197 ELTSPVVAESPKKPQ--VSSLD--EKNSVVDLPCEMKVLSKKNKDKGY----- 245
      :|||||
Qy 241 ETLASIDELATKAIGKKIQONGGL-AVEAGHNGTLA-----GATIS 283
      :|||||
Db 246 -DLATVTKLELKTSDPNKNGSVLEGVKADKSKVKLTISDGLQTTLEVFKEDEKTLVS 304
      :|||||
Qy 284 KLITOKLDGLKNSKLEKINAKKCEDEFTKLEGEHAQ-LGIENVTDENAKKAI 338
      :|||||
Db 305 KKVTSK-DKSTEEKFEKEG-----VSEKIITRADGTRELYTGKSGKAKEVVL 355
      :|||||

```

```

RESULT 2
US-08-158-353-3
; Sequence 3, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cario, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UCT93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-158-353-3

```

```

Query Match 51.1%; Score 945; DB 1; Length 210;
Best Local Similarity 97.9%; Pred. No. 2.3e-64;
Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MACNNSGKDGNTSANSADSVYKGNLTETISKRITDSNAVLLAVKEVALLSSIDEIAKA 60
      :|||||
Db 17 ISCNNSGKDGNTSANSADSVYKGNLTETISKRITDSNAVLLAVKEVALLSSIDEIAKA 76
      :|||||

```

```

Qy 61 ICKIHHNNGLDTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 120
      :|||||
Db 77 ICKIHHNNGLDTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAAKCSPTFN 136
      :|||||
Qy 121 KLEKHTDGLKEGVTDADAKAEALILKANGTKTGAEELGKLFESVEVLSAAKEMLANSVK 180
      :|||||
Db 137 KLEKHTDGLKEGVTDADAKAEALILKANGTKTGAEELGKLFESVEVLSAAKEMLANSVK 196
      :|||||
Qy 181 ELTSPVVAESPKKP 194
      :|||||
Db 197 ELTSPVVAESPKKP 210
      :|||||
RESULT 3
US-08-209-603E-15
; Sequence 15, Application US/08209603E
; Patent No. 6248538
; GENERAL INFORMATION:
; APPLICANT: FUCHS, RENATE
; APPLICANT: WILSKE, BETTINA
; APPLICANT: PREAC-MORSIC, VERA
; APPLICANT: MOTZ, MANFRED
; APPLICANT: SOUTSCHECK, ERWIN
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
; TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCIT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,603E
; FILING DATE: 10-MAR-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP90/02282
; FILING DATE: 21-DEC-1990
; APPLICATION NUMBER: US 07/862,535
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9217-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PROTEIN
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: N/A
; ORIGINAL SOURCE:
; ORGANISM: B. BURGDORFERI
; IMMEDIATE SOURCE:
; LIBRARY: DSM 5662
; POSITION IN GENOME: N/A
; FEATURE:
; IDENTIFICATION METHOD: amino acid analysis
; PUBLICATION INFORMATION: N/A

```

US-08-209-603E-15

Query Match 50.7%; Score 938; DB 4; Length 210;
Best Local Similarity 97.4%; Pred. No. 7.9e-64;
Matches 189; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSIDEIATAA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVLEVEALLSIDEIATAA 76
QY 61 ICKKHONNGLDTEENNHNSSLGAGAYAISTLIKOKLDGKLNKGKIDAAKCCSEFTFN 120
DB 77 ICKKHONNGLDTEENNHNSSLGAGAYAISTLIKOKLDGKLNKGKIDAAKCCSEFTFN 136
QY 121 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 180
DB 137 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 196
QY 181 ELTSPVVAESPCKP 194
DB 197 ELTSPVVAESPCKP 210

RESULT 4

US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Moltz, Manfred
APPLICANT: Sautscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293
CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 209
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match 50.4%; Score 933; DB 4; Length 209;
Best Local Similarity 97.9%; Pred. No. 1.9e-63;
Matches 188; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSIDEIATAA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVLEVEALLSIDEIATAA 76
QY 61 ICKKHONNGLDTEENNHNSSLGAGAYAISTLIKOKLDGKLNKGKIDAAKCCSEFTFN 120
DB 77 ICKKHONNGLDTEENNHNSSLGAGAYAISTLIKOKLDGKLNKGKIDAAKCCSEFTFN 136
QY 121 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 180
DB 137 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 196

QY 181 ELTSPVVAESPCK 192
DB 197 ELTSPVVAESPCK 208

RESULT 5

US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Liff, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Compising
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ. ID NO. 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-30

Query Match 50.4%; Score 932; DB 4; Length 210;
Best Local Similarity 96.9%; Pred. No. 2.2e-63;
Matches 188; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGPNTLEINKKTTDSNAVLAVKEVEALLSIDEIATAA 60
DB 17 ISCNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVLEVEALLSIDEIATAA 76
QY 61 ICKKHONNGLDTEENNHNSSLGAGAYAISTLIKOKLDGKLNKGKIDAAKCCSEFTFN 120
DB 77 ICKKHONNGLDTEENNHNSSLGAGAYAISTLIKOKLDGKLNKGKIDAAKCCSEFTFN 136
QY 121 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 180
DB 137 KLEKHTDGLGEGVTDADAKAELILKANGTKTKGAELGKLFESVEVLSKAAKEMLANSVK 196
QY 181 ELTSPVVAESPCKP 194
DB 197 ELTSPVVAESPCKP 210

RESULT 6

US-08-235-836C-122
; Sequence 122, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ. ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-122

Query Match 50.2%; Score 929; DB 4; Length 588;
Best Local Similarity 89.7%; Pred. No. 1.5e-62;
Matches 191; Conservative 6; Mismatches 16; Indels 0; Gaps 0;
4 NNSGKDGNTSANSADSVKGPNTETINKITDSNAVLAVKEVEALLSSIDEIAKAIGK 63
|||||
DB 286 NNSGKDGNTSANSADSVKGPNTETISKITDSNAVLAVKEVEALLSSIDEIAKAIGK 345
|||||
QY 64 KIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNGELKEKIDAAKCCSEFTTNK 123
|||||
DB 346 KIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNGELKEKIDAAKCCSEFTTNK 405
|||||
QY 124 EKHTDLGKEGYVDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 183
|||||
DB 406 EKHTDLGKEGYVDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 465
|||||
QY 184 SPVVAESPKKPSMVNNSGKDGNTSANSADSVK 216
|||||
DB 466 SPVVAESPKKPGTMAQYVNMHMLNKSASQNVK 498
|||||

RESULT 7
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising

;; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
;; NUMBER OF SEQUENCES: 144
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Brookhaven National Laboratory
;; STREET:
;; CITY: Upton
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 11973
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/235,836C
;; FILING DATE: 29-APR-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/148,191
;; FILING DATE: 01-11-93
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bogosian, Margaret C.
;; REGISTRATION NUMBER: 25,324
;; REFERENCE/DOCKET NUMBER: BNL93-28A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 282-7338
;; TELEFAX: (516) 282-3729
;; INFORMATION FOR SEQ. ID NO: 107:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 466 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-235-836C-107

Query Match 49.5%; Score 915; DB 4; Length 466;
Best Local Similarity 97.4%; Pred. No. 1.3e-61;
Matches 186; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
4 NNSGKDGNTSANSADSVKGPNTETINKITDSNAVLAVKEVEALLSSIDEIAKAIGK 63
|||||
DB 276 NNSGKDGNTSANSADSVKGPNTETISKITDSNAVLAVKEVEALLSSIDEIAKAIGK 335
|||||
QY 64 KIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNGELKEKIDAAKCCSEFTTNK 123
|||||
DB 336 KIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGKNGELKEKIDAAKCCSEFTTNK 395
|||||
QY 124 EKHTDLGKEGYVDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 183
|||||
DB 396 AKHTDLGKEGYVDADAKKAILKANGTKTGABELGKLFESVEVLSKAKEMLANSVKELT 455
|||||
QY 184 SPVVAESPKK 194
|||||
DB 456 SPVVAESPKK 466
|||||

RESULT 8
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620662
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millida Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA


```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US 07/903,580
3      FILING DATE: 25-JUN-1992
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER:  US 07/824,161
6      FILING DATE: 22-JAN-1992
7      PRIOR APPLICATION DATA:
8      APPLICATION NUMBER:  US 07/727,245
9      FILING DATE: 11-JUL-1991
10     ATTORNEY/AGENT INFORMATION:
11     NAME:  BENT, Stephen A.
12     REGISTRATION NUMBER:  29,768
13     REFERENCE/DOCKET NUMBER:  30472/142  IMMD
14     TELECOMMUNICATION INFORMATION:
15     TELEPHONE:  (202)672-5300
16     TELEFAX:  (202)672-5399
17     TELEX:  904136
18     INFORMATION FOR SEQ ID NO: 2:
19     SEQUENCE CHARACTERISTICS:
20     LENGTH: 212 amino acids
21     TYPE: AMINO ACID
22     TOPOLOGY: linear
23     MOLECULE TYPE: protein
24     US-08-031-295-2

```

Query Match	36.2%	Score 670	DB 1	Length 212
Best Local Similarity	73.0%	Pred. No. 1	4e-43	
Matches 143	Conservative 15	Mismatches 36	Indels 2	Gaps 2

QY	MACNNSGKDGNT-SANSDESVKGNGTNEINKRIDSNANVLAVNEVALLSDEIAK	59
Dd	ISCANNSGGGGSASINPAPDESAKGNLTETSKRTIDSNAPFLAKEVEVELYSIDELATK	76
QY	AIGKRHHONNGJDTENNHNHSSLLAGAYAITSLIKOKLDGLANEQ-LKEKIDAKKCSETF	118
Dd	AIGKKIOQNNGILGANADKNSSLLAGAYAITSLITEKITALKNKSGELAKIEDAKCSDEF	136
QY	TNNLKEERTDLCBEVYTDAADAEALKLANKGRTKKAEBELGKFEEVEVLSTRAAKEMLANS	178
Dd	TJKTLAAGAHOAGIDODATNDNSKEALLKTNGTKTAKEBELVKLSEEVASLSKNAOEAASNS	196
QY	VKELTSPVVAESPCKP	194
Dd	VKELTSPVVAETPPKP	212

RESULT 10
 US-07-903-580-2
 : Sequence 2, Application US/07903580
 : Patent No. 6221363
 : GENERAL INFORMATION:
 : APPLICANT: LIVER, Ian
 : APPLICANT: DORNER, Friedrich
 : TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
 : TITLE OF INVENTION: PREVENTION OF LYME DISEASE
 : NUMBER OF SEQUENCES: 3
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Foley & Lardner
 : STREET: 1800 Diagonal Road, Suite 500
 : CITY: Alexandria
 : STATE: Virginia
 : COUNTRY: USA
 : ZIP: 22313-0299
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/903,580
 : FILING DATE: 19920625
 : CLASSIFICATION: 435

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-861-6240
? TELEFAX: 617-861-9540
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 212 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-158-353-4

Query Match          34.1%; Score 631; DB 1; Length 212;
Best Local Similarity 69.9%; Pred. NO. 1,3e-40;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2

Qy      1  MACNNSGRKGNF-SANSADESYKGNLLEINKKTTDSNAVLLAVKEVEALLSIDETAAK 59
Db      17  ISCNNSGKGDSTNPADSEAKGNLLEISKKTTDSNAVFLAVKEVEETLVLSIDETAAK 76

```

[illegible]

```

2
US-09-196-293-11
: Sequence 11, Application US/09196293
: Patent No. 6183755
: GENERAL INFORMATION:
: APPLICANT: Fuchs, Renate
: APPLICANT: Motz, Manfred
: APPLICANT: Soutscheck, Erwin
: APPLICANT: Wilske, Bettina
: APPLICANT: Preacs-Musics, Vera
: TITLE OF INVENTION: Active proteins from Borrelia
: TITLE OF INVENTION: burgdorferi
: FILE REFERENCE: 738.001052
: CURRENT APPLICATION NUMBER: US/09/196,293
: CURRENT FILING DATE: 1998-11-19
: EARLIER APPLICATION NUMBER: US 08/209,603
: EARLIER FILING DATE: 1994-03-10
: EARLIER APPLICATION NUMBER: US 07/862,535
: EARLIER FILING DATE: 1992-06-19
: EARLIER APPLICATION NUMBER: NO PCT/EP90/02282
: EARLIER FILING DATE: 1990-12-21
: EARLIER APPLICATION NUMBER: DE P39 42 728.5
: EARLIER FILING DATE: 1989-12-22
: EARLIER APPLICATION NUMBER: DE P40 18 988.0
: EARLIER FILING DATE: 1990-06-13
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FaastSD for Windows Version 4.0
: SEQ ID NO: 11
: LENGTH: 212
: TYPE: PRT
: ORGANISM: Borrelia burgdorferi
: US-09-196-293-11

```

Query Match	34.0%	Score 629,	DB 4,	Length 212,
Best Local Similarity	70.4%	Pred. NO. 1,8e-40,		
Matches 18;	Conservative 13;	Mismatches 43;	Indels 2;	Gaps 2;
1 MACNNGKDG-NTSANSADSEYKGNLTLEINKTKTDSNAVLAVKEVEALLSTIDEFLAAK 59				

DB 17 ISCNNSGKVGILSTPNPADESAGPVLTEISKITDSNMFVLAKEVEETLVLSIDELAKK 76
QY 60 AIGKRIHONNGLDTENNHNGSLAGAVAI STLTKOKLDGLKN-EGLEKIDAAKCSSEF 118
DB 77 AIGKIDNNNGGLAALNNQNSLAGAVAI STLTEKLSKLNLEELKTEIAKAKCSSEF 136
QY 119 TNKLEKHTDLGKGVTDADAKAELIKANGTKTGAEELKLFESVEVLSKAEMLAN 178
DB 137 TNKLSGHDGLGKODATDHAKAAILKTHATTDKGAKEFKDLFESVEGLKAAQVALTNS 196
QY 179 VKELTSPVAESPCKP 194
DB 197 VKELTSPVAESPCKP 212

RESULT 13
US-08-209-603E-11
; Sequence 11, Application US/08209603E
; Patent No. 6248538

GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSK, BETTINA
APPLICANT: PREAC-MURSC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDOFFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992

ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 357-3355
TELEFAX: (212) 357-5635

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 212
TYPE: AMINO ACID
TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDOFFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:

IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE:
US-08-209-603E-11

Query Match 34.0%; Score 629; DB 4; Length 212;
Best Local Similarity 70.4%; Pred. No. 1.8e-40;
Matches 138; Conservative 13; Mismatches 43; Indels 2; Gaps 2;

QY 1 MACNNSGKDG-NTSANGADSVKGPNTLEINKTKTDSNAVLAVKEVEALLSIDETIAK 59
DB 17 ISCNNSGKVGILSTPNPADESAGPVLTEISKITDSNMFVLAKEVEETLVLSIDELAKK 76
QY 60 AIGKRIHONNGLDTENNHNGSLAGAVAI STLTKOKLDGLKN-EGLEKIDAAKCSSEF 118
DB 77 AIGKIDNNNGGLAALNNQNSLAGAVAI STLTEKLSKLNLEELKTEIAKAKCSSEF 136
QY 119 TNKLEKHTDLGKGVTDADAKAELIKANGTKTGAEELKLFESVEVLSKAEMLAN 178
DB 137 TNKLSGHDGLGKODATDHAKAAILKTHATTDKGAKEFKDLFESVEGLKAAQVALTNS 196
QY 179 VKELTSPVAESPCKP 194
DB 197 VKELTSPVAESPCKP 212

RESULT 14
US-08-235-836C-34
; Sequence 34, Application US/08235836C
; Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93

ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-235-836C-34

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:43 ; Search time 621.2 Seconds

(without alignments)
168.507 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850
Sequence: 1 MACNSGKDGNTSANSADSE.....AVENLAKRAKEMLANSVKEL 377

Scoring table:

BIOSUM62
Gapol 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	377	19	US-09-596-746-30
2	1850	100.0	377	19	US-09-596-746A-30
3	1825	98.6	400	19	US-09-596-746-82
4	1825	98.6	401	19	US-09-596-746A-82
5	1801.5	97.4	374	19	US-09-596-746A-26
6	1796.5	97.1	373	19	US-09-596-746-26
7	1793.5	96.9	397	19	US-09-596-746-62
8	1793.5	96.9	398	19	US-09-596-746A-62
9	1614.5	87.3	378	19	US-09-596-746-28

10	1614.5	87.3	378	19	US-09-596-746A-28	Sequence 28, Appl
11	1589.5	85.9	401	19	US-09-596-746-80	Sequence 80, Appl
12	1589.5	85.9	402	19	US-09-596-746A-80	Sequence 80, Appl
13	1566	84.6	375	19	US-09-596-746A-24	Sequence 24, Appl
14	1561	84.4	374	19	US-09-596-746-24	Sequence 24, Appl
15	1558	84.2	398	19	US-09-596-746-60	Sequence 60, Appl
16	1558	84.2	399	19	US-09-596-746A-60	Sequence 60, Appl
17	1524.5	82.4	370	19	US-09-596-746A-36	Sequence 36, Appl
18	1519.5	82.1	369	19	US-09-596-746-36	Sequence 36, Appl
19	1516.5	82.0	363	19	US-09-596-746-66	Sequence 66, Appl
20	1516.5	82.0	394	19	US-09-596-746A-66	Sequence 66, Appl
21	1485.5	80.3	368	19	US-09-596-746A-42	Sequence 42, Appl
22	1480.5	80.0	367	19	US-09-596-746-42	Sequence 42, Appl
23	1477.5	79.9	381	19	US-09-596-746-74	Sequence 74, Appl
24	1477.5	79.9	382	19	US-09-596-746A-74	Sequence 74, Appl
25	1472	79.6	385	19	US-09-596-746A-56	Sequence 56, Appl
26	1467	79.3	384	19	US-09-596-746-56	Sequence 56, Appl
27	1465	79.2	560	19	US-09-596-746A-52	Sequence 52, Appl
28	1465	79.2	560	19	US-09-596-746A-52	Sequence 52, Appl
29	1464	79.1	408	19	US-09-596-746-78	Sequence 78, Appl
30	1464	79.1	409	19	US-09-596-746A-78	Sequence 78, Appl
31	1289	69.4	371	19	US-09-596-746A-34	Sequence 34, Appl
32	1284	69.7	370	19	US-09-596-746-34	Sequence 34, Appl
33	1281	69.2	394	19	US-09-596-746-64	Sequence 64, Appl
34	1281	69.2	395	19	US-09-596-746A-64	Sequence 64, Appl
35	1264.5	68.4	370	19	US-09-596-746A-32	Sequence 32, Appl
36	1259.5	68.1	369	19	US-09-596-746-32	Sequence 32, Appl
37	1256.5	67.9	363	19	US-09-596-746-68	Sequence 68, Appl
38	1256.5	67.9	369	19	US-09-596-746A-68	Sequence 68, Appl
39	1250	67.6	369	19	US-09-596-746A-40	Sequence 40, Appl
40	1245	67.3	368	19	US-09-596-746-40	Sequence 40, Appl
41	1242	67.1	392	19	US-09-596-746-72	Sequence 72, Appl
42	1242	67.1	393	19	US-09-596-746A-72	Sequence 72, Appl
43	1225.5	66.2	368	19	US-09-596-746-38	Sequence 38, Appl
44	1225.5	66.2	368	19	US-09-596-746A-38	Sequence 38, Appl
45	1217.5	65.8	391	19	US-09-596-746-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-09-596-746-30
Sequence 30, Application US/09596746
GENERAL INFORMATION:
APPLICANT: Datwyler, Raymond J.
APPLICANT: Seinstost, Gerald
APPLICANT: Dykhuisen, Daniel
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 377
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Ospec Chimera
US-09-596-746-30

Query Match 100.0%; Score 1850; DB 19; Length 377;
Best local similarity 100.0%; Pred. No. 3.4e-125;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 MACNSGKDGNTSANSADSEYKGNLFIKKITDSNVLAVKEVALLSSIDEIAKA 60

```
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEINKTIDTSNAVLLAVEEALLSIDEIAKA 60
Qy 61 IGKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Db 61 IGKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Qy 121 KLEKHTDGLKEGYTDADAKAELIKANGTKGAEEGLKFESVEVLSKAAKEMLANSVK 180
Db 121 KLEKHTDGLKEGYTDADAKAELIKANGTKGAEEGLKFESVEVLSKAAKEMLANSVK 180
Qy 181 ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 240
Db 181 ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 240
Qy 241 ETLASIDELATRAIGKTIQONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSKELK 300
Db 241 ETLASIDELATRAIGKTIQONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSKELK 300
Qy 301 EKIEKAKCSSEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAEELEKFKAVE 360
Db 301 EKIEKAKCSSEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAEELEKFKAVE 360
Qy 361 NIKAKEMLANSVKEL 377
Db 361 NIKAKEMLANSVKEL 377

RESULT 2
US-09-596-746a-30
; Sequence 30, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-30
```

```
Query Match 100.0%; Score 1850; DB 19; Length 377;
Best Local Similarity 100.0%; Pred. No. 3.4e-125;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACNNSGKDGNTSANSADSVKGPMLTEINKTIDTSNAVLLAVEEALLSIDEIAKA 60
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEINKTIDTSNAVLLAVEEALLSIDEIAKA 60
Qy 61 IGKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Db 61 IGKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 120
Qy 121 KLEKHTDGLKEGYTDADAKAELIKANGTKGAEEGLKFESVEVLSKAAKEMLANSVK 180
Db 121 KLEKHTDGLKEGYTDADAKAELIKANGTKGAEEGLKFESVEVLSKAAKEMLANSVK 180
Qy 181 ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 240
Db 181 ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 240
```

```
Db 181 ELTSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 240
Qy 241 ETLASIDELATRAIGKTIQONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSKELK 300
Db 241 ETLASIDELATRAIGKTIQONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSKELK 300
Qy 301 EKIEKAKCSSEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAEELEKFKAVE 360
Db 301 EKIEKAKCSSEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAEELEKFKAVE 360
Qy 361 NIKAKEMLANSVKEL 377
Db 361 NIKAKEMLANSVKEL 377

RESULT 3
US-09-596-746-82
; Sequence 82, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 400
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-82
```

```
Query Match 98.6%; Score 1825; DB 19; Length 400;
Best Local Similarity 98.9%; Pred. No. 2.4e-123;
Matches 372; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ACNNSGKDGNTSANSADSVKGPMLTEINKTIDTSNAVLLAVEEALLSIDEIAKAI 61
Db 25 SCNNSGKDGNTSANSADSVKGPMLTEISKITIDTSNAVLLAVEEALLSIDEIAKAI 84
Qy 62 GKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 121
Db 85 GKTHONNGDTEENNNGSLAGAVAIISLTIKOKLDGLKNEGKEKIDAAKCSSEFTN 144
Qy 122 LKEKHTDGLKEGYTDADAKAELIKANGTKGAEEGLKFESVEVLSKAAKEMLANSVK 181
Db 145 LKEKHTDGLKEGYTDADAKAELIKANGTKGAEEGLKFESVEVLSKAAKEMLANSVK 204
Qy 182 LSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 241
Db 205 LSPVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITTSNAVLAVERI 264
Qy 242 ETLASIDELATRAIGKTIQONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSKELK 301
Db 265 ETLASIDELATRAIGKTIQONGGLAVEAGHNGTLAGAVTISKITOKLDGLKNSKELK 324
Qy 302 KIEKAKCSSEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAEELEKFKAVE 361
Db 325 KIEKAKCSSEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAEELEKFKAVE 384
Qy 362 LAKAKEMLANSVKEL 377
Db 385 LAKAKEMLANSVKEL 400
```



```
RESULT 4
; US-09-596-746a-82
; Sequence 82, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIORITY FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 84
; TYPE: PRT
; ORGANISM: OspC Chimera
; LENGTH: 401
; -09-596-746a-82

Query Match          98.6%; Score 1825; DB 19; Length 401;
Best Local Similarity 98.9%; Pred. No. 2,4e-123;
Matches 37/2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSVKGNLTPEINKITDSNAVLLAVEVEALLSSIDEIAAKAI 61
DB 26 SCNNSKDGNTSANSADSVKGNLTPEISKITDSNAVLLAVEVEALLSSIDEIAAKAI 85
OY 62 GKRIHONNGIDTENNHNHSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTNK 121
DB 86 GKRIHONNGIDTENNHNHSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTNK 145
OY 122 LKEKHHDLCGEVTDADAKAAILKANGTKTKGAEBELGKLFESVEVLSKAKEMLANSVKE 181
DB 146 LKEKHHDLCGEVTDADAKAAILKANGTKTKGAEBELGKLFESVEVLSKAKEMLANSVKE 205
OY 182 LTPSPVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTPEISKITDSNAVLLAVEVEI 241
DB 206 LTPSPVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTPEISKITDSNAVLLAVEVEI 265
OY 242 TLLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVYITSLITQKLDGLKNSERLKE 301
DB 266 TLLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVYITSLITQKLDGLKNSERLKE 325
OY 302 KIENAKKCSDEFTKLEGEHAQGIENVTDENAKKAILITDAKDKGAELEKLEKFAVEN 361
DB 326 KIENAKKCSDEFTKLEGEHAQGIENVTDENAKKAILITDAKDKGAELEKLEKFAVEN 385
OY 362 LAKAKEMLANSVKE 377
DB 386 LAKAKEMLANSVKE 401

RESULT 5
; US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIORITY FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
```

```
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
; US-09-596-746a-26

Query Match          97.4%; Score 1801.5; DB 19; Length 374;
Best Local Similarity 98.1%; Pred. No. 1.1e-121;
Matches 37/0; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

OY 1 MACNNSGKDGNTSANSADSVKGNLTPEINKITDSNAVLLAVEVEALLSSIDEIAAKA 60
DB 1 MACNNSGKDGNTSANSADSVKGNLTPEISKITDSNAVLLAVEVEALLSSIDEIAAKA 60
OY 61 GKRIHONNGIDTENNHNHSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTN 120
DB 61 GKRIHONNGIDTENNHNHSLAGAVAI STLKQKLDGLKNEGLKEKIDAAKCSSEFTN 120
OY 121 KLEKHHDLCGEVTDADAKAAILKANGTKTKGAEBELGKLFESVEVLSKAKEMLANSVKE 180
DB 121 KLEKHHDLCGEVTDADAKAAILKANGTKTKGAEBELGKLFESVEVLSKAKEMLANSVKE 180
OY 181 ELTPSPVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTPEISKITDSNAVLLAVEVEI 240
DB 181 ELTPSPVAESPKRPSMVNNSGKDGNTSANSADSVKGNLTPEISKITDSNAVLLAVEVEI 237
OY 241 ETLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVYITSLITQKLDGLKNSERLKE 300
DB 238 ETLASIDELATKAIGKKIQONGSLAVEAGHNGTLLAGAVYITSLITQKLDGLKNSERLKE 297
OY 301 KIENAKKCSDEFTKLEGEHAQGIENVTDENAKKAILITDAKDKGAELEKLEKFAVE 360
DB 298 KIENAKKCSDEFTKLEGEHAQGIENVTDENAKKAILITDAKDKGAELEKLEKFAVE 357
OY 361 NLAKAKEMLANSVKE 377
DB 358 NLAKAKEMLANSVKE 374

RESULT 6
; US-09-596-746-26
; Sequence 26, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIORITY FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
; US-09-596-746-26

Query Match          97.1%; Score 1796.5; DB 19; Length 373;
```

Best Local Similarity 98.1%; Pred. No. 2,4e-121;
Matches 369; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINIKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 60
OY 62 GKTIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGLKNEGJKEKIDAKKCEFTTNK 121
DB 61 GKTIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGLKNEGJKEKIDAKKCEFTTNK 120
OY 122 LKEKHTDGLKEGYTDADAEALILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 181
DB 121 LKEKHTDGLKEGYTDADAEALILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 180
OY 182 LTPVVAESPKRPSMVNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 241
DB 181 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 237
OY 242 TLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTSKITLTKDGLKNSKEK 301
DB 238 TLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTSKITLTKDGLKNSKEK 297
OY 302 KIENAKKCEDEFTKKLEGEHAOLGIENVTDENAKKAILITDAADKGAELKFKAVEN 361
DB 298 KIENAKKCEDEFTKKLEGEHAOLGIENVTDENAKKAILITDAADKGAELKFKAVEN 357
OY 362 LAKAKEMLANSVKEL 377
DB 358 LAKAKEMLANSVKEL 373
```

RESULT 7

```
US-09-596-746-62
; Sequence 62, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoist, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 397
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-62
```

Query Match 96.9%; Score 1793.5; DB 19; Length 397;
Best Local Similarity 97.9%; Pred. No. 4,4e-121;
Matches 368; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINIKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 25 SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 84
OY 62 GKTIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGLKNEGJKEKIDAKKCEFTTNK 121
DB 85 GKTIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGLKNEGJKEKIDAKKCEFTTNK 144
OY 122 LKEKHTDGLKEGYTDADAEALILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 181
DB 145 LKEKHTDGLKEGYTDADAEALILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 204
```

```
OY 182 LTPVVAESPKRPSMVNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 241
DB 205 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 261
OY 242 TLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTSKITLTKDGLKNSKEK 301
DB 262 TLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTSKITLTKDGLKNSKEK 321
OY 302 KIENAKKCEDEFTKKLEGEHAOLGIENVTDENAKKAILITDAADKGAELKFKAVEN 361
DB 322 KIENAKKCEDEFTKKLEGEHAOLGIENVTDENAKKAILITDAADKGAELKFKAVEN 381
OY 362 LAKAKEMLANSVKEL 377
DB 382 LAKAKEMLANSVKEL 397
```

RESULT 8

```
US-09-596-746a-62
; Sequence 62, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoist, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746a-62
```

Query Match 96.9%; Score 1793.5; DB 19; Length 398;
Best Local Similarity 97.9%; Pred. No. 4,4e-121;
Matches 368; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINIKKTTDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 85
OY 62 GKTIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGLKNEGJKEKIDAKKCEFTTNK 121
DB 86 GKTIHONNGLDTEYNNHNSLAGAVALSTLIKOKLDGLKNEGJKEKIDAKKCEFTTNK 145
OY 122 LKEKHTDGLKEGYTDADAEALILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 181
DB 146 LKEKHTDGLKEGYTDADAEALILKANGTKGAEELGKLFESVEVLSKAKEMLANSYKE 205
OY 182 LTPVVAESPKRPSMVNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 241
DB 206 LTPVVAES---PAMVNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 262
OY 242 TLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTSKITLTKDGLKNSKEK 301
DB 263 TLLASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTTSKITLTKDGLKNSKEK 322
OY 302 KIENAKKCEDEFTKKLEGEHAOLGIENVTDENAKKAILITDAADKGAELKFKAVEN 361
DB 323 KIENAKKCEDEFTKKLEGEHAOLGIENVTDENAKKAILITDAADKGAELKFKAVEN 382
OY 362 LAKAKEMLANSVKEL 377
DB 383 LAKAKEMLANSVKEL 398
```

```
RESULT 9
US-09-596-746-28
; Sequence 28, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-28

Query Match      87.3%; Score 1614.5; DB 19; Length 378;
Best Local Similarity 87.8%; Pred. No. 3.6e-108;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
QY 61 ICKKIHNNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
DB 61 ICKKIHNNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
QY 121 KLEKHTDGLKEGVTDAKAEAILKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDAKAEAILKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
QY 121 KLEKHTDGLKEGVTDAKAEAILKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDAKAEAILKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
DB 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
QY 241 ETLTASIDELATKAIGKKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
DB 241 ETLTASIDELATKAIGKKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
QY 301 EKIEENKCKSEDTTKLEGEHAQGIENVTDENAKKAILITDAKDKGALEKLFKAVE 360
DB 301 EKIEENKCKSEDTTKLEGEHAQGIENVTDENAKKAILITDAKDKGALEKLFKAVE 360
QY 360 AETEKAKCKSEPTAKLKGHTDLGKEGVTDDNAKKAAILTNDKTGGADELEKLFESVK 359
DB 360 AETEKAKCKSEPTAKLKGHTDLGKEGVTDDNAKKAAILTNDKTGGADELEKLFESVK 359

RESULT 10
US-09-596-746A-28
; Sequence 28, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
```

```
FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-28

Query Match      87.3%; Score 1614.5; DB 19; Length 378;
Best Local Similarity 87.8%; Pred. No. 3.6e-108;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
DB 1 MACNSGKDGNTSANSADSVKGNLTETINKKITDSNAVLAVKEVEALLSSIDEIAKA 60
QY 61 ICKKIHNNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
DB 61 ICKKIHNNGDPTENNHNHSLAGAYASTLIKOKLDGLKNEGLKEKIDAKKCEFTTN 120
QY 121 KLEKHTDGLKEGVTDAKAEAILKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 121 KLEKHTDGLKEGVTDAKAEAILKANGTKTGAEELGKLFESVEVLSKAKEMLANSVK 180
QY 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
DB 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADSVKGNLTETISKITTESNAVLAVKEI 240
QY 241 ETLTASIDELATKAIGKKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
DB 241 ETLTASIDELATKAIGKKIQONGGLAVAGHNGTLLAGATTISKLITOKLDGLKNSKELK 300
QY 301 EKIEENKCKSEDTTKLEGEHAQGIENVTDENAKKAILITDAKDKGALEKLFKAVE 360
DB 301 EKIEENKCKSEDTTKLEGEHAQGIENVTDENAKKAILITDAKDKGALEKLFKAVE 360
QY 360 AETEKAKCKSEPTAKLKGHTDLGKEGVTDDNAKKAAILTNDKTGGADELEKLFESVK 359
DB 360 AETEKAKCKSEPTAKLKGHTDLGKEGVTDDNAKKAAILTNDKTGGADELEKLFESVK 359

RESULT 11
US-09-596-746-80
; Sequence 80, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRF
; ORGANISM: OspC Chimera
US-09-596-746-80
```

Query Match 85.9%; Score 1589.5; DB 19; Length 401;
 Best Local Similarity 86.7%; Pred. No. 2.5e-106;
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEVALLSIDETIAAKAI 61
 DB 25 SCNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEVALLSIDETIAAKAI 84
 QY 62 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 121
 DB 85 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 144
 QY 122 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVKE 181
 DB 145 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVKE 204
 QY 182 LTPVVAESPCKPSPMVNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEV 241
 DB 205 LTPVVAESPCKPSPMVNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEV 264
 QY 242 TLASIDELATKAIGKIQONGGLAVEAGHNGTLAGAYTTSKLTOKLOGKSEKKE 301
 DB 265 TLTSIDELA-KAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISAKIDSGELKA 323
 QY 302 KIENAKKCEDEFTKKEGEHQAOLGIENVTDENAKKAILITDAKDGALEKLEKRAVEN 361
 DB 324 EIEAKKCESEFTAKLGEHTDGLKEGVTDNNAKAILITNDKTGGADELKLEKESVK 383
 QY 362 LAKAAKEMLANSVKEL 377
 DB 384 LSKAAKEMLTNSVKEL 399

RESULT 12

US-09-596-746a-80
 ; Sequence 80, Application US/09596746A
 ; GENERAL INFORMATION:
 ; APPLICANT: Dattwyler, Raymond J.
 ; APPLICANT: Seinst, Gerald
 ; APPLICANT: Dykhuisen, Daniai
 ; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746A
 ; PRIOR FILING DATE: 2000-06-19
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 80
 ; LENGTH: 402
 ; TYPE: PRT
 ; ORGANISM: ospC Chimera
 ; US-09-596-746a-80

Query Match 85.9%; Score 1589.5; DB 19; Length 402;
 Best Local Similarity 86.7%; Pred. No. 2.5e-106;
 Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

QY 2 ACNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEVALLSIDETIAAKAI 61
 DB 26 SCNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEVALLSIDETIAAKAI 85
 QY 62 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 121
 DB 86 GKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTNK 145
 QY 122 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLKESVYVLSKAAKEMLANSVKE 181

DB 146 LKEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVYVLSKAAKEMLANSVKE 205
 QY 182 LTPVVAESPCKPSPMVNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEV 241
 DB 206 LTPVVAESPCKPSPMVNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEV 265
 QY 242 TLASIDELATKAIGKIQONGGLAVEAGHNGTLAGAYTTSKLTOKLOGKSEKKE 301
 DB 266 TLTSIDELA-KAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISAKIDSGELKA 324
 QY 302 KIENAKKCEDEFTKKEGEHQAOLGIENVTDENAKKAILITDAKDGALEKLEKRAVEN 361
 DB 325 EIEAKKCESEFTAKLGEHTDGLKEGVTDNNAKAILITNDKTGGADELKLEKESVK 384
 QY 362 LAKAAKEMLANSVKEL 377
 DB 385 LSKAAKEMLTNSVKEL 400

RESULT 13

US-09-596-746a-24
 ; Sequence 24, Application US/09596746A
 ; GENERAL INFORMATION:
 ; APPLICANT: Dattwyler, Raymond J.
 ; APPLICANT: Seinst, Gerald
 ; APPLICANT: Dykhuisen, Daniai
 ; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746A
 ; PRIOR FILING DATE: 2000-06-19
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24
 ; LENGTH: 375
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: OspC Chimera
 ; US-09-596-746a-24

Query Match 84.8%; Score 1566; DB 19; Length 375;
 Best Local Similarity 85.9%; Pred. No. 1.1e-104;
 Matches 324; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

QY 1 MACNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEVALLSIDETIAAKAI 60
 DB 1 MACNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEVALLSIDETIAAKAI 60
 QY 61 IGKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTN 120
 DB 61 IGKRIHONNGLDTEYNNHNSLAGAAYASTLIKOKLDGKNEGLKEKIDAAKCEFTTN 120
 QY 121 KLEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVK 180
 DB 121 KLEKHTDGLKEGVTDADAKAELILKANGTKYKGAELGLFESVEVLSKAAKEMLANSVK 180
 QY 181 ELTPVVAESPCKPSPMVNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEV 240
 DB 181 ELTPVVAESPCKPSPMVNNSGKDNTSANSADSVKGNLTETIKKTTSDNAVLAVKEV 237
 QY 241 ETLASIDELATKAIGKIQONGGLAVEAGHNGTLAGAYTTSKLTOKLOGKSEKKE 300
 DB 238 ETLTSIDELA-KAIGKRIKNDVSLDNEADHNGSLISGAYLSNLTITKISAKIDSGELK 296
 QY 301 EKIENAKKCEDEFTKKEGEHQAOLGIENVTDENAKKAILITDAKDGALEKLEKRAVE 360
 DB 297 AEIEAKKCESEFTAKLGEHTDGLKEGVTDNNAKAILITNDKTGGADELKLEKESVK 356

OY 361 NLAKAKEMLANSVKEL 377
 ||:|||||
 Db 357 NLSKAKEMLTNSVKEL 373

RESULT 14 US-09-596-746-24

; Sequence 24, Application US/09596746
 ; GENERAL INFORMATION:
 ; APPLICANT: Datwyler, Raymond J.
 ; APPLICANT: Sehnost, Gerald
 ; APPLICANT: Dykhutzen, Daniel
 ; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746
 ; CURRENT FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; PRIOR FILING DATE: 1999-06-18
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24
 ; LENGTH: 374
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: OspC Chimera
 US-09-596-746-24

Query Match 84.4%; Score 1561; DB 19; Length 374;
 Best Local Similarity 85.9%; Pred. No. 2,6e-104;
 Matches 323; Conservative 22; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
 ||:|||||
 Db 1 ACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVEALLSSIDEIAAKAI 60
 OY 62 GKTIHONGIDTEYNNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAAKCCSEPTNK 121
 ||:|||||
 Db 61 GKTIHONGIDTEYNNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAAKCCSEPTNK 120
 OY 122 LKKEHTDLGEGYTDADAKAAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSVKE 181
 ||:|||||
 Db 121 LKKEHTDLGEGYTDADAKAAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSVKE 180
 OY 182 LTPVVAESPKPSMVNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVE 241
 ||:|||||
 Db 181 LTPVVAES---PAMVNNSSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVE 237
 OY 242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNEGLKE 301
 ||:|||||
 Db 238 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKA 296
 OY 302 KIENAKKCEDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVEN 361
 ||:|||||
 Db 297 EIEKAKKCEDEFTAKLGEHTDLGEGYTDONAKKAILITNDKTGGADELKLFESVKN 356
 OY 362 LAKAKEMLANSVKEL 377
 ||:|||||
 Db 357 LSKAKEMLTNSVKEL 372

RESULT 15
 US-09-596-746-60
 ; Sequence 60, Application US/09596746
 ; GENERAL INFORMATION:
 ; APPLICANT: Datwyler, Raymond J.
 ; APPLICANT: Sehnost, Gerald
 ; APPLICANT: Dykhutzen, Daniel

; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746
 ; CURRENT FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; PRIOR FILING DATE: 1999-06-18
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 60
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: OspC Chimera
 US-09-596-746-60

Query Match 84.2%; Score 1558; DB 19; Length 398;
 Best Local Similarity 85.6%; Pred. No. 4.7e-104;
 Matches 322; Conservative 23; Mismatches 27; Indels 4; Gaps 2;

OY 2 ACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
 ||:|||||
 Db 25 ACNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVEALLSSIDEIAAKAI 84
 OY 62 GKTIHONGIDTEYNNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAAKCCSEPTNK 121
 ||:|||||
 Db 85 GKTIHONGIDTEYNNHNGSLAGAYASTLIKOKLDGKNEGLKEKIDAAKCCSEPTNK 144
 OY 122 LKKEHTDLGEGYTDADAKAAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSVKE 181
 ||:|||||
 Db 145 LKKEHTDLGEGYTDADAKAAILKANGTKTKGAELGKLFESVEVLSKAKEMLANSVKE 204
 OY 182 LTPVVAESPKPSMVNNSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVE 241
 ||:|||||
 Db 205 LTPVVAES---PAMVNNSSGKDGNTSANSADSVKGPMLTEINKITDSNAVLAVKEVE 261
 OY 242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNEGLKE 301
 ||:|||||
 Db 262 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKA 320
 OY 302 KIENAKKCEDEFTKLEGEHAOLGIENVTDENAKKAILITDAKDKGAELKLFKAVEN 361
 ||:|||||
 Db 321 EIEKAKKCEDEFTAKLGEHTDLGEGYTDONAKKAILITNDKTGGADELKLFESVKN 380
 OY 362 LAKAKEMLANSVKEL 377
 ||:|||||
 Db 381 LSKAKEMLTNSVKEL 396

Search completed: March 18, 2002, 10:08:44
 Job time: 972 sec


```

; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSpC-derived peptide fragments
; FILE REFERENCE: 459-666p
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-974-992-7

```

```

Query Match          34.1%; Score 631; DB 6; Length 212;
Best Local Similarity 69.9%; Pred. No. 1e-38;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY 1 MACNSGKQNT-SANSADSVKGNLFEINKRTDSNAVLAVEVALLSIDETIAAK 59
Db 17 ISCNNGKGGDSASTNPADSAKGNLFEISKRTDSNAFLAVEVETLSTIDELAKK 76
QY 60 AIGKTIHONNGLDTENNHNSGLAGAVASTLIKOKLDGLN-EGLEKIDAAKCSFT 118
Db 77 AIGOKIDNNNGLAALNNNGSLAGAVASTLIKOKLDGLN-EGLEKIDAAKCSFT 136
QY 119 TNKLEKHTDGLKEGVTADAKAELKANGTKTGAEELGLFESVVLAKAKEMLAN 178
Db 137 TTKLKSGRHDLGKQDADTDHAKKAILKTHATTDKGAKEFKDLFESVGLAKAVALTNS 196
QY 179 VKELTSPVAESPKRP 194
Db 197 VKELTSPVAESPKRP 212

```

```

RESULT 3
US-09-974-992-3
Sequence 3, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSpC-derived peptide fragments
FILE REFERENCE: 459-666p
CURRENT APPLICATION NUMBER: US/09/974,992
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-974-992-3

```

```

Query Match          32.2%; Score 596.5; DB 6; Length 207;
Best Local Similarity 66.0%; Pred. No. 3.2e-36;
Matches 128; Conservative 25; Mismatches 38; Indels 3; Gaps 2;

```

```

QY 1 MACNSGKQNT-SANSADSVKGNLFEINKRTDSNAVLAVEVALLSIDETIAAK 60
Db 17 ISCNNGKGGDSASTNPADSAKGNLFEISKRTDSNAFLAVEVETLSTIDELAKK 76
QY 61 IGGKTIHONNGLDTENNHNSGLAGAVASTLIKOKLDGLN-EGLEKIDAAKCSFT 120
Db 74 IGGKTIHONNGLDTENNHNSGLAGAVASTLIKOKLDGLN-EGLEKIDAAKCSFT 133
QY 121 KLEKHTDGLKEGVTADAKAELKANGTKTGAEELGLFESVVLAKAKEMLAN 180
Db 134 KLEKHTDGLKEGVTADAKAELKANGTKTGAEELGLFESVVLAKAKEMLAN 193
QY 181 ELTSPVAESPKRP 194
Db 194 ELTSPVAESPKRP 207

```

```

RESULT 4
US-09-708-427-19883
Sequence 19883, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
PRIOR FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..1014
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc-feature
LOCATION: 1..1014
OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

```

```

Query Match          9.3%; Score 172.5; DB 6; Length 1014;
Best Local Similarity 23.2%; Pred. No. 0.00018;
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

QY 8 KQNTSANSADSVKGNLFEINKRTDSNAVLAVEVALLSIDETIAAK 45
Db 401 KHEGTEADSKGYLGVAELHSTLEAFQVYKSSLEALNATENEKELEINAAVTSEKK 460
QY 46 VEALESIDEIETIAAK-----AIGKTIHONNGLDTENNHNSGLAGAVASTLIKOKLDGLN-EGLEKIDAAKCSFT 120
Db 461 LEA-----TVEYEVKISSEBENLESTRNELNVTGQ-KLESTINDKAGLDESEVMEKLS 516
QY 95 KLDGLKNSGLKEKIDAA-----KKCSFTFNK-----LKEK 126
Db 517 ABESELEQK--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRSEASSLLEK 574
QY 127 TDLEKGEVTDADAKAELKANGTKTGAEELGLFESVVLAKAKEMLANVKEL 182
Db 575 RDLBEKIKSYEQLAEASGKSSLEKLEQTLGRLAAEVSNEKLEQEPDQAKESLOS 634
QY 183 -TSPVAESPKRP-----PSMNVNSGKQNTSANSADSVKGNLFEISKRTDSNA 233
Db 635 SESELEAFTNNOLEKIKIQLEGLIGSGSVKETAALRLERALE-----RPNKETSDDL 689
QY 234 VLAKEIETILASIDELATAIG-----KKIQONG-----GLAVEA 269
Db 690 VEKTKHTENOIEEYKKAHAEASGVADTRKVELEDALSKLNESITIEELGACQGLERES 749
QY 270 GHNGTL-----LAGAVYTSKLITQKLDGLK-NSKLEKELEMAKCSDEFTKRL-EG 319

```


Db 750 GDIAEYVNLKLNLELANHGSANELQTKLSALEAEKQFANLEASKTIEDTLTKQLTSEG 809
QY 320 EHAQIGIENVTDENAKKALITDAADKGAAELEK-----LFRAVENLAKAAEM-- 369
Db 810 EKIQSOIEKLRVAAAEKSVL-----ESHFELEKTLSEVKAQKLEVENAATAVSVAE 863
QY 370 LANSYKE 376
Db 864 LITSKLOE 870

RESULT 5

US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19882
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1018
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1018
; OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19882

Query Match 9.3%; Score 172.5; DB 6; Length 1018;
Best Local Similarity 23.2%; Pred. No. 0.00018;
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

QY 8 KDGNTSANSAD-----SVKGPNL-----TEINKKITDS-NAVLLAVKE 45
Db 405 KHGETEADSKGYIGVAVELQSTLEAFQVXSSLSLEALNITATENEKELTENLNAVTSKKK 464
QY 46 VEALLSIDELIAK-----AIGKKIHONGLDTENNHNGLAGAVAI STL--IKQ 94
Db 465 LEA---TVDEYSVKISESENLESIRNELNVTQG-KLESTENDKKAAGLQSEVMEKLS 520
QY 95 KLDGLKNEGLKEKIDAA-----KCSFTPTNK-----LKEKH 126
Db 521 AEESLEQKG--REIDEATTKRMELALHOSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 578
QY 127 TDL-GREGYTDADAKAILKANGTKTKGAELGKLFESVYVLSAKAKEMLANSVKEL-- 182
Db 579 RDEGKIKSYEEOGLAEGSSSLKEKLEQTLGRLAAESVNEKLEKQEPDOAEKSLQSS 638
QY 183 -TSPVVAESPCK-----PSMVNNSGKDGNTSANSADSVKGPNLTEISKITTESNAV 233
Db 639 SESELLAETNNOLKIKITQIEGLEIGSGSVKEKRLAKLEAIE-----RNQKTESSDL 633
QY 234 VLAVKEITLLASIDELATKAIG-----KKIQONG---GLAVEA 269
Db 694 VEKLKTHENQIEEYKRLAHASGVADTRKVELDALSKLKNLESTIEELGAKCGLEKES 733
QY 270 GHNGTL-----LAGAYITSKLITQKLDGLK-NSEKLIKIEIAKACSDPTFKL--EG 319
Db 754 GDIAEYVNLKLNLELANHGSANELQTKLSALEAEKQFANLEASKTIEDTLTKQLTSEG 813
QY 320 EHAQIGIENVTDENAKKALITDAADKGAAELEK-----LFRAVENLAKAAEM-- 369
Db 814 EKIQSOIEKLRVAAAEKSVL-----ESHFELEKTLSEVKAQKLEVENAATAVSVAE 867

QY 370 LANSYKE 376
Db 868 LITSKLOE 874

RESULT 6

US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881

Query Match 9.3%; Score 172.5; DB 6; Length 1269;
Best Local Similarity 23.2%; Pred. No. 0.00024;
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

QY 8 KDGNTSANSAD-----SVKGPNL-----TEINKKITDS-NAVLLAVKE 45
Db 656 KHGETEADSKGYIGVAVELQSTLEAFQVXSSLSLEALNITATENEKELTENLNAVTSKKK 715
QY 46 VEALLSIDELIAK-----AIGKKIHONGLDTENNHNGLAGAVAI STL--IKQ 94
Db 716 LEA---TVDEYSVKISESENLESIRNELNVTQG-KLESTENDKKAAGLQSEVMEKLS 771
QY 95 KLDGLKNEGLKEKIDAA-----KCSFTPTNK-----LKEKH 126
Db 772 AEESLEQKG--REIDEATTKRMELALHOSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 829
QY 127 TDL-GREGYTDADAKAILKANGTKTKGAELGKLFESVYVLSAKAKEMLANSVKEL-- 182
Db 830 RDEGKIKSYEEOGLAEGSSSLKEKLEQTLGRLAAESVNEKLEKQEPDOAEKSLQSS 889
QY 183 -TSPVVAESPCK-----PSMVNNSGKDGNTSANSADSVKGPNLTEISKITTESNAV 233
Db 890 SESELLAETNNOLKIKITQIEGLEIGSGSVKEKRLAKLEAIE-----RNQKTESSDL 944
QY 234 VLAVKEITLLASIDELATKAIG-----KKIQONG---GLAVEA 269
Db 945 VEKLKTHENQIEEYKRLAHASGVADTRKVELDALSKLKNLESTIEELGAKCGLEKES 1004
QY 270 GHNGTL-----LAGAYITSKLITQKLDGLK-NSEKLIKIEIAKACSDPTFKL--EG 319
Db 1005 GDIAEYVNLKLNLELANHGSANELQTKLSALEAEKQFANLEASKTIEDTLTKQLTSEG 1064
QY 320 EHAQIGIENVTDENAKKALITDAADKGAAELEK-----LFRAVENLAKAAEM-- 369
Db 1065 EKIQSOIEKLRVAAAEKSVL-----ESHFELEKTLSEVKAQKLEVENAATAVSVAE 1118
QY 370 LANSYKE 376
Db 1119 LITSKLOE 1125

RESULT 7

```

US-09-742-096-5
; Sequence 5, Application US/09742096
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 2007/3050DIY
; CURRENT APPLICATION NUMBER: US/09/742, 096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-5

```

```

Query Match          9.1%; Score 168; DB 6; Length 630;
Best Local Similarity 18.9%; Pred. No. 0.0002;
Matches 82; Conservative 104; Mismatches 144; Indels 104; Gaps 20;

```

```

QY 9 DGNTS-----ANSADSVK-----GNLTKKIDSNVAVLAVKVEEALLSID 54
DB 135 DGSAASSVEERSIASVDESIDSIEENVAPTEVEI---VAPTEVEIAPSVESVAP 191
QY 55 ELAKAIGKIKHONGDTEENNINGSILAGAVASTLIKOKLDGLKNEGKIDAKK 114
DB 192 ESVAAPSVESVAEN-----VEESVAENVEEIVAPSVESV- AENV 230
QY 115 SEFTNKIKKEKHTDKEGVTADAKAEILKANGTKTGAEL- GKLEESV-----EVL 167
DB 231 EESVAENVEESVAENVEESVAE- NVEESV-----AENVVEIAPTEVESVAPTEVEI 282
QY 168 SKAKKEMLANSVKELTSPVVAESPKRPSVNNNGKDGNTS-ANSADSVKGNLTK-ISK 225
DB 283 APVSESVAETVEEIVAPSVES- VAPSVESVAENVEESVAENVEESV- AENVESVAE 340
QY 226 KITESNAVAVLAVEIEFLA-SIDELATKAIGKIKOONGGLAVEGHNGTLLAGAVT--I 282
DB 341 NVEESVA-----ENVEEIVAPSVETVEIAPTEVESVAENVATNLSNLLSGIETE 395
QY 283 SKLITOKLDGLKNS-----EKLK-----EKINAKKSEDFTKLEGEHA 322
DB 396 KOSILNEIEVEKENVTTILEKEVETTAESVTTFSNILEIEIDENTITNDTIEKLELHE 455
QY 323 QI-----GIENVTDENAKKAL-----ITDAKDKCAAELEKLFKAVENLA 363
DB 456 NVLSALETTOSEBEKEKVIDIEVEKEVATTLTIEVQAESEESTITEIFENLIEBNA 515
QY 364 KAKKEMLANSVKEL 377
DB 516 VESNEKVAENLEKL 529

```

```

RESULT 8
US-09-742-096-3
; Sequence 3, Application US/09742096
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 2007/3050DIY
; CURRENT APPLICATION NUMBER: US/09/742, 096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06

```

```

; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

```

```

Query Match          8.6%; Score 160; DB 6; Length 1786;
Best Local Similarity 18.7%; Pred. No. 0.0031;
Matches 76; Conservative 97; Mismatches 167; Indels 66; Gaps 16;

```

```

QY 19 ESKGPNLTKKIDSNVAVLAVKVEEALLSIDELAKAIGKIKHONGDTEENNHN 78
DB 588 EEIVAPTEVEI---VAPTEVEIAPSVESVAPSVESVEEENVEESVAENVEESVAENVE 644
QY 79 GSLAG-----AYAISTLIKOKLDGLKNEGKIDAKKSEFTNKLKKEKHTDGLK 133
DB 645 ESVAENVEEIVAPTEVEIAPTEVEIAPSVESV- APSVEESVEEENVEESVAENVEES 702
QY 134 VTDADAKKAILKANGTKTGAELGKLEESVYLSKAKKEMLANSVKELTSPVVAES--P 191
DB 703 VAE- NVEESV- AENVESVAENV-----BEIVAPTEVEIAPTEVEIAPSVESVAP 753
QY 192 KPSVNNNGKDGNTSANSADSVKGNLTK-ISKKITESNAVAVLAVEI-----ETIL 244
DB 754 SVEESVEEENVEE--SVANVEESV- AENVESVAENVEES- VAPTEVEIAPSVESVA 808
QY 245 ASIDELATKAIGKIKOON-----GLAVEGHNGTLLAGAVTISKLITOKLDGLKNS-- 296
DB 809 PSVESVAENVATNLSNLLSGIETEIRKISILNEIEVEKENVTTILEVEETTA 868
QY 297 -----EKLKIKINAKKSEDFTKLEGEHAOL---GIENVTDENAKKAL-----ETIL 339
DB 869 ESVTTFSNILEIEIDENTITNDTIEKLELHENVLSALETTOSEBEKEKVIDIEVEKE 928
QY 340 -----ITDAKDKCAAELEKLFKAVENLANAKKEMLANSVKEL 377
DB 929 EVATTLTIEVQAEKSNANTITEIFENLIEBNAVESNENVAENLEKL 974

```

```

RESULT 9
US-09-614-150-6213
; Sequence 6213, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637

```

PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 6213
LENGTH: 1639
TYPE: PRP
ORGANISM: DROSOPHILA
US-09-614-150-6213

Query Match 8.6%; Score 159.5; DB 6; Length 1639;
Best Local Similarity 23.9%; Pred. No. 0.003;
Matches 105; Conservative 58; Mismatches 162; Indels 115; Gaps 19;

QY 2 ACNSGKDCNTSANSAD-----ESYKGPLTEINKITDSNAVLAKEVEALLSIDE 55
DB 1107 ARNSGDCGQTYAEVIDDLHKHLDSVR-EHLVSADKFOADANGEIRARQNTYILDQITE 1165
QY 56 IAKAIGKKIHQNGLDTENNHNSLLAGAVASTLIKOKLDGLKNEGLKEKIDAAKKS 115
DB 1166 NAKKEL-----QALDLNDEGQALARAKKESVEFGQ-----SEQISDISREA 1210
QY 116 EFTFNKLEK-EKHTDLGEGVTDAAKEALILKAN-----GFKTKG--AEELGK 1159
DB 1211 BALADLSEADQDLKNA-----KDAKDAVEKAHQALAKSAIDLOKIGTELSEVGLSH 1266
QY 160 LRESVEVLSKAKE-----MLANSYKELTSPVVAESPCKPSNVNNSGKDGNT 206
DB 1267 VKOSLCTVQTSKEALRKANEVYDTALTLLNDVNRQTOPEIDISQK-----KDAVA 1318
QY 207 SANSADSVKGPLTEISKITTESNAVLAKEIETLLASIDELATKAIGKKIQ----- 261
DB 1319 ANERADELTK--QITELS-----NSNGELFPADEFTQEOL-----TEALLKRAEQOOLE 1365
QY 262 -----NGLAVEAGHNGTLGAVYTSKLITOKLDGLKNSK-----KL 299
DB 1366 IELLERAKAHDKATKAVEGDN-TLKEANNTEYKELAGPQSDVQORSESEKALQYTPNI 1424
QY 300 KEKIEKAKCSEDEFTKKLEGEHAQDGIENVTDENAKKAIL--TTDAKDKGALEKLFK 357
DB 1425 EKEIQVAESLISQAEALDGAANKN--ANBAKKNQAEQALKYAEQASKD-----AELLR 1475
QY 358 AVENLAKAKEMLANSVKEL 377
DB 1476 RKANETKVAARNLREBADDL 1495

RESULT 10
US-09-708-427-3730
Sequence 3730, Application US/09708427
GENERAL INFORMATION:

APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 3730
LENGTH: 805
TYPE: PRP
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..805
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..805
OTHER INFORMATION: Ceres Seq. ID 1810877
US-09-708-427-3730

Query Match 8.5%; Score 156.5; DB 6; Length 805;
Best Local Similarity 20.4%; Pred. No. 0.0019;
Matches 86; Conservative 83; Mismatches 167; Indels 85; Gaps 16;

QY 21 VKGPNLTEINK-----KITDSNAVLAKE--VEAL-----LSSIDEIAKAIKGIH 66
DB 50 VKHKELEVEKAEADGLGLELNSRKMIIEEDRIRISALAEKLELQKOSASLEEKLK 109
QY 67 QNNGLDTENNHNSLLAGAVASTLIKOKLDGLKNEGLKEK-----DAKKS 115
DB 110 IS--DERYSKTALLSQALSONSVLEOKIKSL--BELSEKVSPLKSAIIVAEDEGKKS 164
QY 116 ---ETFNKLEKHTDGRGVTDAKKE--AIILKA-----NGKKTGAELGKLF 162
DB 165 IOMOEYOEKSKLESSLNQSSARNSLEEDLRALQGAHEDIGNVSTRSVELQGLFQ 224
QY 163 SVEVLKSAKEMT-----ANSYKELTSPVVAESPCKPS-----WVNSGKDGNT 206
DB 225 TSQKLEKAEKLEKDLAIVQKNSLEATLSVAMEKERDLSENLNAMKELKSEERLEK 284
QY 207 SANSADSVK-----GPNLTEISKITTESNAVLAKEIETLLASIDEL 250
DB 285 QAREIDEATRTSIELEALHKHSELKYQKTWEDESSRDTFAKSLTEKSKDEEKIRVEYGK 344
QY 251 ATKAIKGIKIQNGGLAVEAGHNGTLGAVYTSKLITOKLDGLKNSK--LKEKIEKAK 308
DB 345 LAERCGQSILQELDQSSAEN-ELLADTNQKIKIQLEGLGYDSKEFAIEKLNOKPT 403
QY 309 CSEDEFTKKLEGEHAQDGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAKE 368
DB 404 EANDLITKLSH-----ENVIEEH-KROYLEASGVADTKVEVEEALLNLNLTSESTIE 456
QY 369 M 369
DB 457 L 457

RESULT 11
US-09-708-427-3729
Sequence 3729, Application US/09708427
GENERAL INFORMATION:

APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 3729
LENGTH: 841
TYPE: PRP
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..841
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..841
OTHER INFORMATION: Ceres Seq. ID 1810876
US-09-708-427-3729

Query Match 8.5%; Score 156.5; DB 6; Length 841;
Best Local Similarity 20.4%; Pred. No. 0.002;
Matches 86; Conservative 83; Mismatches 167; Indels 85; Gaps 16;

QY 21 VKGPNLTEINK-----KITDSNAVLAKE--VEAL-----LSSIDEIAKAIKGIH 66
DB 86 VKHKELEVEKAEADGLGLELNSRKMIIEEDRIRISALAEKLELQKOSASLEEKLK 145
QY 67 QNNGLDTENNHNSLLAGAVASTLIKOKLDGLKNEGLKEK-----DAKKS 115

```

Db 146 IS---DERYKTDALLSQNSVLEQKLKSL--EELSEKVELSKALIYAEBGRKSS 200
QY 116 ---ETFTNKLKEKHTDLKEGVTADAKE-----ALKA-----NGRTKGAELGLIFE 162
Db 201 IOMQYOEKVELSLSSLSMOSSARNSLEEDLRIALQKGAHEHDIGNVSKRSVLEGLIFQ 260
QY 163 SVEVLAKAKEML-----ANSVKELTSPVAESPSPKPS-----AVNNSGKQNT 206
Db 261 TSQLEKAEKELKDLAIOVKNSSLEATLSVAMEKEHDLSENLNAVMEKKSSEERLEK 320
QY 207 SANSDESVK-----GPNLTETSKKTTESNAVVLAVKEITLLASIDEL 250
Db 321 QAREIDEFTTTSIELEALHKHSEKLVOKTMEFSSRDTFAKSTKESKDLKIRVEYEGK 380
QY 251 ATKALGKIOQNGILAVAGHNGTLLAGAVTISKLTIOKLDGLKNSRK--LKEKIEKAK 308
Db 381 LAECGGSLSLQELDSSAEN--ELADTNNQKIKIQLEGLYDSKEFAIEKLNOKDT 439
QY 309 CSEDTKLEGEHAGLQIENVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 368
Db 440 EAKDLITKLKSH-----ENVIEEH--KROYLEASGVADTRKVEVEALLKNTLESTIEE 492
QY 369 M 369
Db 493 L 493

```

RESULT 12

```

US-09-708-427-32624
; Sequence 32624, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32624
; LENGTH: 1038
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1038
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1038
; OTHER INFORMATION: Ceres Seq. ID 1834662
US-09-708-427-32624

```

```

Query Match 8.4%; Score 155.5; DB 6; Length 1038;
Best Local Similarity 25.0%; Pred. No. 0.0032;
Matches 104; Conservative 53; Mismatches 186; Indels 73; Gaps 16;

```

```

QY 7 GKDQNTSAN---SADESXKGPMLTEINKKTTDSNAVLAVKEVEALLSSIDEIAAKAIK 63
Db 372 GKEGNEFCNVISDAEKSQSVINIPVANHLOEPRN---IPVKLSEHLPKPTPEPT-----K 423
QY 64 KIHONNGIDTENNNHNGSLAGA-----YAISTLIKOKLDGLKNEGKLEKIDAK--KC 114
Db 424 RIAKNEPVKSTYKKEQSSSSSEASKIPVCLRVDPLEPRNGSKSVSHPRMKSKEKTI 483
QY 115 SETFTNKLKEKHTDLGKESGVTADAKKAILKANGTKTKGAELGLKLFESVVLAKAKEM 174
Db 484 AAPSSKKAESRTYPEACNVKCEDANEMKMAEGSLNALRTKNG-----SVESNSNLQDES 539
QY 175 LANSVKELTSPVAESPSPKPSMVNNSGKDGNTSANSDESVKGPNTL--EISKITESNA 232
Db 540 NGELIKPEAKENREQPAKKSFTEE-----AARIIQSMYKRYDVRWPEIKKLKEIAT 593

```

```

QY 233 VLAIV-----KEIETLLASIDELATKAIGKIOQNGL-----AVEAGHNG--TLIA 277
Db 594 VREOMGDVKKRIBALEASTDQHIEE--KEIYVNGSLVNNLLKIDAVGLHPSTREFPK 650
QY 278 GATTISKLTIOKLDGLKNS-----EKLKEKIEKAKCSEDTFKLEGEHAGLQIEN--V 329
Db 651 ALATELSSIODKIDSLKNSCASAKEKVAKEVOYEIKSQPSD---SPVNLHQSOLTEENKMY 707
QY 330 TDENAKKAILI-----TDAKDKGAELKLFKAVENLAKAKAKEMLANS 373
Db 708 SDTNLEKVLRLSPDEHPMSVLRNRTDEKQASAEETEGYGLFETLATDSKQATENA 763

```

RESULT 13

```

US-09-708-427-32623
; Sequence 32623, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32623
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1042
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1042
; OTHER INFORMATION: Ceres Seq. ID 1834661
US-09-708-427-32623

```

```

Query Match 8.4%; Score 155.5; DB 6; Length 1042;
Best Local Similarity 25.0%; Pred. No. 0.0032;
Matches 104; Conservative 53; Mismatches 186; Indels 73; Gaps 16;

```

```

QY 7 GKDQNTSAN---SADESXKGPMLTEINKKTTDSNAVLAVKEVEALLSSIDEIAAKAIK 63
Db 376 GKEGNEFCNVISDAEKSQSVINIPVANHLOEPRN---IPVKLSEHLPKPTPEPT-----K 427
QY 64 KIHONNGIDTENNNHNGSLAGA-----YAISTLIKOKLDGLKNEGKLEKIDAK--KC 114
Db 428 RIAKNEPVKSTYKKEQSSSSSEASKIPVCLRVDPLEPRNGSKSVSHPRMKSKEKTI 487
QY 115 SETFTNKLKEKHTDLGKESGVTADAKKAILKANGTKTKGAELGLKLFESVVLAKAKEM 174
Db 488 AAPSSKKAESRTYPEACNVKCEDANEMKMAEGSLNALRTKNG-----SVESNSNLQDES 543
QY 175 LANSVKELTSPVAESPSPKPSMVNNSGKDGNTSANSDESVKGPNTL--EISKITESNA 232
Db 544 NGELIKPEAKENREQPAKKSFTEE-----AARIIQSMYKRYDVRWPEIKKLKEIAT 597
QY 233 VLAIV-----KEIETLLASIDELATKAIGKIOQNGL-----AVEAGHNG--TLIA 277
Db 598 VREOMGDVKKRIBALEASTDQHIEE--KEIYVNGSLVNNLLKIDAVGLHPSTREFPK 654
QY 278 GATTISKLTIOKLDGLKNS-----EKLKEKIEKAKCSEDTFKLEGEHAGLQIEN--V 329
Db 655 ALATELSSIODKIDSLKNSCASAKEKVAKEVOYEIKSQPSD---SPVNLHQSOLTEENKMY 711
QY 330 TDENAKKAILI-----TDAKDKGAELKLFKAVENLAKAKAKEMLANS 373
Db 712 SDTNLEKVLRLSPDEHPMSVLRNRTDEKQASAEETEGYGLFETLATDSKQATENA 767

```

RESULT 14

```
US-09-708-427-32622
; Sequence 32622, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32622
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1043
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1043
; OTHER INFORMATION: Ceres Seq. ID 1834660
; -09-708-427-32622

Query Match      8.4%; Score 155.5; DB 6; Length 1043;
Best Local Similarity 25.0%; Pred. No. 0.0032;
Matches 104; Conservative 53; Mismatches 186; Indels 73; Gaps 16;

QY 7 GKDQNTSAN---SADESVGPNLTETINKKITDSDNAVLAVKEVEALLSIDETIAAIGK 63
DB 377 GKEGNECVLSAEKSSVINIPVANHLOEPN---IPKLSENHLPRTEPT-----K 428
QY 64 KIHQNGLDTENNHNSLLAGA-----YAISTLIKOKLDGLKNGLEKIDAK--KC 114
DB 429 RIAKNPVASTKKEQSSSEASKLPVCLRVDPPLPERNGSKSVSHPRMEKSETKI 488
QY 115 SEFTTKLEKHTDLEKEGYTDADAKKALIKANGTKKGAEEGLKLPESVEVLSKAKE 174
DB 489 AAPLSKKKESRTVPACVNCEDANAEMKMGSLNALRTEKG---SVESNSNLQEES 544
QY 175 LANSVVELSPVVAESPKKPSMVNNSGKDQNTSANSADSVKPNLT--EISKITESNA 232
DB 545 NGETIIPCEAKKREOPAKKSTPEE-----AARIIQSMYRCYDVRMRPEIKKLEIAT 598
QY 233 VYLA---KEIETLASIDELATKAIGKKIQONGI-----AVEAGHG--TLA 277
DB 599 VREQMGDVKKRIETALASTDQHIE--KEIYVNGELVMNLLKLDVAVESGLHPSIREFR 655
QY 278 GAYTISKLTOKLDGLKNS-----EKLKEIKINAKKCEDEFTKKLEGEHAQIGEN--V 329
DB 656 ALATELSSYDQKLDLSLKNCSASAKEAVKEQVIRKQSPSD--SPVNLHESQLTEENK 712
QY 330 TDENAKKAILI-----TDAKDKGALELEKLFKAVENLAKAKEMLAN 373
DB 713 SDTNLEKVLRLSPPEHPMSVLANRTDEKQASAEETEBGIGLFTLATDSQATENA 768

RESULT 15
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jlang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7646
```

```
LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: 1881..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

Query Match      8.3%; Score 153; DB 7; Length 1881;
Best Local Similarity 23.8%; Pred. No. 0.011;
Matches 91; Conservative 59; Mismatches 146; Indels 86; Gaps 16;

QY 27 TEINKKITDSDNAVLAVKEVEALLSIDETIAAIGKIH--QNGLDTENNHNSLLAG 84
DB 1529 TELKDRISE-----VEKRAMISENSEFVIREKSDIKLSIKINSKESKE--- 1577
QY 85 AYASTLIKOKLDGLKNE--GLKEKIDAAKCCSETFTNKLKEKHTDLEKRGVTDADAK 142
DB 1578 ---ITTHNEOK-TSLKODIAKLSQDHESAOQTOLEKENOLKELKASLEKHNTESA 1632
QY 143 ILKANGTKTGAEELGKLPESVEVLSKAKEMLANSYKELTSPVVAESPKKPSMVNSG 202
DB 1633 -----EKNQIKELSETISLTKTELKTSGDALKQSOKE-----KITLTKNSDT 1676
QY 203 DG-----MTSANSADSVKGNLTETISKITITESNAVLAVKEIETL---LAS 246
DB 1677 ESKLEKQLEELKVKSDLOQADEKLKGITREIATLK-----SELETYKNSGLST 1725
QY 247 IDELA--TKAIGKKIQONGLAVEAGHNGTLAGAVYISKLTOKLDGLKN--SEKLEK 302
DB 1726 TSELAALTFTKSLKEKEKEELOPLSGKSKLEBEDYQKHSIDISKALALDELKEXTK 1785
QY 303 IENAKKCE--DFT--KKLEGEHAQIG--IENVTDENAKKAILITPAARDKGAEELEK 355
DB 1786 DSKKKTLELNDLTSTKKLEFTEKQTSKFNLEER-----KDKELVKLNKE 1833
QY 356 FKAVENLAKAKEMLANVKEL 377
DB 1834 LELLKNDNSGAKKELSEKVKSL 1855
```

Search completed: March 18, 2002, 09:58:21
Job time: 379 sec

Tue Mar 19 10:57:48 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:47 ; Search time 68.77 Seconds
(without alignments)
417.592 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850

Sequence: 1 MACNNNSGRKDGTSANSADES.....AVENIAKAKEMLANSVKEL 377

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_68:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	51.1	210	2	G70218
2	938	50.7	210	2	S69927
3	876	47.4	193	2	S70287
4	863	46.6	191	2	I40153
5	853	46.1	177	2	I40129
6	826	44.6	178	2	I40125
7	823	44.5	178	2	I40104
8	774.5	41.9	211	2	I40145
9	773.5	41.8	193	2	S70279
10	739	39.9	178	2	I40124
11	709	38.3	210	2	I40144
12	695.5	37.6	211	2	S69918
13	692.5	37.4	211	2	I40277
14	690.5	37.3	193	2	S70280
15	686.5	37.1	211	2	I40268
16	686	37.1	190	2	S70273
17	685.5	37.1	209	2	I40273
18	683.5	36.9	191	2	S70278
19	675	36.5	194	2	S70277
20	674.5	36.5	209	2	S69917
21	670	36.2	212	2	S70254
22	666.5	36.0	193	2	S70276
23	664.5	35.9	193	2	S70274
24	664.5	35.9	211	2	I40278
25	664	35.9	212	2	S69922
26	662.5	35.8	191	2	S70284
27	662.5	35.8	193	2	S70265
28	662	35.8	194	2	S70268
29	661.5	35.8	211	2	S69932

30	657.5	35.5	209	2	I40142	outer surface prot
31	655	35.4	212	2	I40079	outer surface prot
32	650	35.1	212	2	I40143	outer surface prot
33	646	34.9	214	2	S69916	outer surface prot
34	641.5	34.7	193	2	S70286	outer surface prot
35	638	34.5	194	2	S70289	outer surface prot
36	635.5	34.4	191	2	S70288	outer surface prot
37	635.5	34.4	203	2	I40108	outer surface prot
38	631	34.1	212	2	S20543	outer surface prot
39	630.5	34.1	209	2	I40281	outer surface prot
40	630.5	34.1	209	2	I40285	outer surface prot
41	630	34.1	192	2	S70285	outer surface prot
42	628	33.9	210	2	S69925	outer surface prot
43	627	33.9	210	2	S69923	outer surface prot
44	627	33.9	212	2	S69921	outer surface prot
45	622.5	33.6	209	2	S69926	outer surface prot

ALIGNMENTS

RESULT 1
G70218
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C:Accession: G70218; I40269; S37726; S70281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; M01D:98065943
A:Accession: G70218
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <KLE>
A:Cross-references: GB:AE000792; NID:93453098; PIDN:AAC66329.1; PID:92689901; TIGR:BB
A:Experimental source: strain B31
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: I40269; M01D:96025162
A:Accession: I40269
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:ID49497; NID:9707092; PIDN:BA08457.1; PID:9769684
R:Jauris-Heipke, S.; Fuchs, R.; Mocz, W.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os
A:Reference number: S37726; M01D:93268136
A:Accession: S37726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <JAU>
A:Cross-references: EMBL:X69596; NID:9311391; PIDN:CAA49306.1; PID:9311392
R:River, I.; Gibbs, C.P.; Schuster, R.; Dornier, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; M01D:96296448
A:Accession: S70281
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210 <LIY>
A:Cross-references: EMBL:I42887; NID:9858715; PIDN:AB36995.1; PID:91695212
A:Experimental source: strain Ip2
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 51.1%; Score 945; DB 2; Length 210;
 Best Local Similarity 97.9%; Pred. No. 5.5e-39;
 Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTNSANSADSVKGPMLTEINKKITSNNAVLAVKEVALLSSIDEIAAKA 60
 DB 17 ISCNNSGKDGTNSANSADSVKGPMLTEISKITDSNAVLAVKEVALLSSIDEIAAKA 76
 QY 61 IGGKIHONNGLDTENNHNSILLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCSSEFTFN 120
 DB 77 IGGKIHONNGLDTENNHNSILLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCSSEFTFN 136
 QY 121 KIKKHHDLGKEGYTDADAKKAILKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVK 180
 DB 137 KIKKHHDLGKEGYTDADAKKAILKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVK 196
 QY 181 ELTSPVAESPCKP 194
 DB 197 ELTSPVAESPCKP 210

RESULT 2
 569927
 outer surface protein C precursor - Lyme disease spirochete (strain PKa)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 A:Variety: strain PKa
 C:Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
 C:Accession: S69927; S72659
 R:Jauris-Heldke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.
 J. Clin. Microbiol. 33, 1860-1866, 1995
 A>Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
 A:Reference number: 140047; MUID:95395018
 A:Accession: S69927
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-210 <JAU>
 A:Cross-references: EMBL:X69589
 R:Jauris, S.
 A:Experimental source: strain PKa
 submitted to the EMBL Data Library, February 1994
 A:Reference number: S72659
 A:Accession: S72659
 A:Molecule type: DNA
 A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAM>
 A:Cross-references: EMBL:X69589
 C:Genetics:
 A:Gene: ospC
 A:Superfamily: Lyme disease spirochete surface protein C

Query Match 50.7%; Score 938; DB 2; Length 210;
 Best Local Similarity 96.9%; Pred. No. 1.2e-38;
 Matches 188; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTNSANSADSVKGPMLTEINKKITSNNAVLAVKEVALLSSIDEIAAKA 60
 DB 17 ISCNNSGKDGTNSANSADSVKGPMLTEISKITDSNAVLAVKEVALLSSIDEIAAKA 76
 QY 61 IGGKIHONNGLDTENNHNSILLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCSSEFTFN 120
 DB 77 IGGKIHONNGLDTENNHNSILLAGAVAI STLKOKLDGLKNEGLKEKIDAAKCSSEFTFN 136
 QY 121 KIKKHHDLGKEGYTDADAKKAILKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVK 180
 DB 137 KIKKHHDLGKEGYTDADAKKAILKANGTKTKGAEELGKLFESVEVLSKAKEMLANSVK 196
 QY 181 ELTSPVAESPCKP 194
 DB 197 ELTSPVAESPCKP 210

RESULT 3
 570287

outer surface protein C - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: S70287
 R:Livley, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
 Mol. Microbiol. 18, 237-269, 1995
 A>Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70287
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <LIV>
 A:Cross-references: EMBL:L42895; NID:9858723; PIDN:AA837003.1; PID:91695220
 A:Experimental source: strain 28354
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 47.4%; Score 876; DB 2; Length 193;
 Best Local Similarity 100.0%; Pred. No. 1e-35;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 NSGKDGTNSANSADSVKGPMLTEISKITDSNAVLAVKEVALLSSIDEIAAKA 61
 DB 2 NSGKDGTNSANSADSVKGPMLTEISKITDSNAVLAVKEVALLSSIDEIAAKA 61
 QY 258 KIQONGGLAVAGNHGTLGAVYISKLITOKLDGLKNSKLEKIEENAKKCSSEFTTKL 317
 DB 62 KIQONGGLAVAGNHGTLGAVYISKLITOKLDGLKNSKLEKIEENAKKCSSEFTTKL 121
 QY 318 GEHQAOLGIENVVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 377
 DB 122 GEHQAOLGIENVVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 181

RESULT 4
 140153
 outer surface protein C - Lyme disease spirochete (fragment)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
 C:Accession: I40153
 R:Funf, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
 Infect. Immun. 62, 3213-3221, 1994
 A>Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi
 A:Reference number: I40153; MUID:94314437
 A:Accession: I40153
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-191 <RES>
 A:Cross-references: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562
 C:Genetics:
 A:Gene: ospC
 C:Superfamily: Lyme disease spirochete surface protein C

Query Match 46.6%; Score 863; DB 2; Length 191;
 Best Local Similarity 99.4%; Pred. No. 4.2e-35;
 Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 NSGKDGTNSANSADSVKGPMLTEISKITDSNAVLAVKEVALLSSIDEIAAKA 60
 DB 1 NSGKDGTNSANSADSVKGPMLTEISKITDSNAVLAVKEVALLSSIDEIAAKA 60
 QY 259 IQONGGLAVAGNHGTLGAVYISKLITOKLDGLKNSKLEKIEENAKKCSSEFTTKLE 318
 DB 61 IQONGGLAVAGNHGTLGAVYISKLITOKLDGLKNSKLEKIEENAKKCSSEFTTKLE 120
 QY 319 GEHQAOLGIENVVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 377
 DB 121 GEHQAOLGIENVVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEMLANSVKEL 179

RESULT 5
140129
outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40129; S54199
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; M0ID:95286481
A:Accession: I40129
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-177 <RES>
A:Cross-references: EMBL:X84783; NID:g793825; PIDN:CAA59254.1; PID:g793826
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 46.1%; Score 853; DB 2; Length 177;
Best Local Similarity 98.3%; Pred. No. 1.1e-34;
Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 11 NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATKAIGKKIQONG 70
|||||
DB 1 NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATKAIGKKIQONG 60
QY 71 LPTENNNGSLAGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 130
|||||
DB 61 LPTENNNGSLAGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 120
QY 131 KEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSVKELTSPV 187
|||||
DB 121 KEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSVKELTSPV 177

RESULT 6
140125
outer surface protein C - Lyme disease spirochete (strain MDL) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40125; S54199
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; M0ID:95286481
A:Accession: I40125
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL:X84779; NID:g793817; PIDN:CAA59250.1; PID:g793818
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 44.6%; Score 826; DB 2; Length 178;
Best Local Similarity 98.8%; Pred. No. 2.3e-33;
Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 205 NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATKAIGKKIQONG 264
|||||
DB 1 NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATKAIGKKIQONG 60
QY 265 LAVEAGHNGTLGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 324
|||||
DB 61 LAVEAGHNGTLGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 120
QY 325 GIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAEMLANSVKEL 377
|||||
DB 121 GIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAEMLANSVKEL 173
RESULT 7

140104
outer surface protein C - Lyme disease spirochete (strain 272) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40104; S54184
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; M0ID:95286481
A:Accession: I40104
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL:X84785; NID:g793787; PIDN:CAA59256.1; PID:g793788
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 44.5%; Score 823; DB 2; Length 178;
Best Local Similarity 98.3%; Pred. No. 3.2e-33;
Matches 170; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 205 NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATKAIGKKIQONG 264
|||||
DB 1 NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATKAIGKKIQONG 60
QY 265 LAVEAGHNGTLGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 324
|||||
DB 61 LAVEAGHNGTLGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 120
QY 325 GIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAEMLANSVKEL 377
|||||
DB 121 GIEVNTDENAKKAILITDAKDKGAELKLFKAVENLAKAEMLANSVKEL 173

RESULT 8
140145
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40145
R:Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A:Title: Expression and sequence of outer surface protein C among North American isol
A:Reference number: I40143; M0ID:95154673
A:Accession: I40145
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 41.9%; Score 774.5; DB 2; Length 211;
Best Local Similarity 81.0%; Pred. No. 8.1e-31;
Matches 158; Conservative 14; Mismatches 22; Indels 1; Gaps 1;
QY 1 MACNNSKDGK-NTSANSADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATK 59
:|||||
DB 17 ISCNNSKDGKMASTNPADESVKGPNTLEIKTKITSDSNVAVLAVKEVETLLASIDELATK 76
QY 60 AIGKKIHONNGCLPTENNNGSLAGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 119
|||||
DB 77 AIGKKIHONNGCLPTENNNGSLAGAVATISLITOKLDGKLNSEKLEKIENAKKCSDFPKKLEGEHAOL 136
QY 120 NKLEKHTDGLKEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSV 179
|||||
DB 137 NKLEKHTDGLKEGVTADAKKAILITDAKDKGAELKLFESVEVLSKAEMLANSV 196
QY 180 KELTSPVAVESPKRP 194

Db 197 KELTSPPVAESPCKP 211

|||||

RESULT 9

S70279

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000

C:Accession: S70279

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A:Reference number: S70255; MUID:96296448

A:Accession: S70279

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-193 <LIV>

A:Cross-references: EMBL:U42898; NID:9858729; PIDN:AB37007.1; PID:91695223

A:Experimental source: strain 25015

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 41.8%; Score 773.5; DB 2; Length 193;

Matches 158; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

QY 3 CNGSGKDGNTSANSADSVKGNLTETINKKIDTSNAVLLAVEVEALLSIDELAKAI 61

Db 1 CNGSGKDGNTSANSADSVKGNLTETINKKIDTSNAVLLAVEVEALLSIDELAKAI 60

QY 62 GKTHQNNGLDTENNHNGSLAGAYASTLIKQKLDGKNEKKEIDAKKCEFTTK 121

Db 61 GKTHQNNGLDTENNHNGSLAGAYASTLIKQKLDGKNEKKEIDAKKCEFTTK 120

QY 122 LKEKHTDGLKRGVTDADAKKAILKANGTKTGAEELGKLFESVEYLSKAEMLANSVKE 181

Db 121 LKSHTELGKQDADODDADKAILRTHTKDKGAELDKLFEPENLSKAEMLANSVKE 180

QY 182 LTPSPVAESPCKP 194

Db 181 LTPSPVAESPCKP 193

QY 10

124

outer surface protein C - Lyme disease spirochete (strain KIPP) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000

C:Accession: I40124; S54194

R:Thelissen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

J. Bacteriol. 177, 3036-3044, 1995

A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A:Reference number: I40104; MUID:95286481

A:Accession: I40124

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178 <RES>

A:Cross-references: EMBL:X84782; NID:9793815; PIDN:CAA59253.1; PID:9793816

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 39.9%; Score 739; DB 2; Length 178;

Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 205 NTSANSADSVKGNLTETINKKIDTSNAVLLAVEVEALLSIDELAKAIKAGKTIQONG 264

Db 1 NTSANSADSVKGNLTETINKKIDTSNAVLLAVEVEALLSIDELAKAIKAGKTIQONG 60

QY 265 LAVEAGHNGTLAGAYTISKLTITQKLDGKNEKKEIDAKKCEFTTKLEGEHAOL 324

|||||

Db 61 LAVEAGHNGTLAGAYTISKLTITQKIRWIENSKIRGIENAKKCEFTTKLEGEHAOL 120

QY 325 GTEVNDENNAKKAILLTDAAKKGAAELEKFAVENLAKAKEMLANSVKEI 377

Db 121 GTEVNDENNAKKAILLTDAAKKGAAELEKFAVENLPOAKEMLANSVKEI 173

RESULT 11

I40144

outer surface protein C precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000

C:Accession: I40144; S70282

R:Stevenson, B.; Barthold, S.W.

FEMS Microbiol. Lett. 124, 367-372, 1994

A:Title: Expression and sequence of outer surface protein C among North American isol

A:Reference number: I40143; MUID:95154673

A:Accession: I40144

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-210 <RES>

A:Cross-references: EMBL:U04281; NID:9434663; PIDN:AA043297.1; PID:9434664

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di

A:Reference number: S70255; MUID:96296448

A:Accession: S70282

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 19-210 <LIV>

A:Cross-references: EMBL:U42893; NID:9858721; PIDN:AB37001.1; PID:91695218

A:Experimental source: strain 297

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 38.3%; Score 709; DB 2; Length 210;

Matches 147; Conservative 22; Mismatches 24; Indels 2; Gaps 2;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETINKKIDTSNAVLLAVEVEALLSIDELAKA 60

Db 17 ISCNNSGKDGNTSANSADSVKGNLTETINKKIDTSNAVLLAVEVEALLSIDELAKA 75

QY 61 IGKTHQNNGLDTENNHNGSLAGAYASTLIKQKLDGKNEKKEIDAKKCEFTT 119

Db 76 IGKTHQNNGLDTENNHNGSLAGAYASTLIKQKLDGKNEKKEIDAKKCEFTT 135

QY 120 NTLKHTDGLKRGVTDADAKKAILKANGTKTGAEELGKLFESVEYLSKAEMLANSV 179

Db 136 AKLKGHTDGLKRGVTDADAKKAILKANGTKTGAEELGKLFESVEYLSKAEMLANSV 195

QY 180 KELTSPPVAESPCKP 194

Db 196 KELTSPPVAESPCKP 210

RESULT 12

S69918

outer surface protein C precursor - Lyme disease spirochete (strain PBre)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

A:Variety: strain PBre

C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000

C:Accession: S69918; S72674; I40103

R:Tauris-Helpe, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutacheck

J. Clin. Microbiol. 33, 1860-1866, 1995

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre

A:Reference number: I40047; MUID:95395018

A:Accession: S69918

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-211 <JAD>
A:Cross-references: EMBL:X81522; NID:g872021; PID:CAA57242.1; PID:g872022
A:Experimental source: strain PBre
R:Roessler, D.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72674
A:Accession: S72674
A:Molecule type: DNA
A:Residues: 1-152, 'E', 154-211 <ROE>
A:Cross-references: EMBL:X81522; NID:g872021; PID:CAA57242.1; PID:g872022
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```

Db      77  ||||| ||| | : ||||| ||| || : | | : | : ||||| : |
      120  NKLKKEHTDLDGEGVTDADAKKAITLAKANGTKRGAEELGKLPESYEVLAKAKEMLANSV 179
      137  KRLSDNOAELGIENKTDNNAKAILKTHNAKDGAELVKLSESVAGLLKAAQAILANSV 196
Oy      180  KELTSPVVAESPCKP 194
      197  KELTSPVVAESPCKP 211
Db

```

Search completed: March 18, 2002, 09:56:47
 Job time: 370 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:45 ; Search time 39.62 Seconds

(without alignments)
348.880 Million cell updates/sec

Title: US-09-596-746A-30

Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES.....AVENIAKAKEMIANSVKEL 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008
Listing first 45 summaries

Database : SWISSprot_39:*

Result No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	51.1	210	1	OSCL_BORBU
2	631	34.1	212	1	OSCL_BORBU
3	397.5	21.5	214	1	VM24_BORHE
4	397	21.5	215	1	VM03_BORHE
5	174.5	9.4	1509	1	MYSN_ACACA
6	172	9.3	492	1	M5_STRPY
7	164.5	8.9	1164	1	BAG_STRPY
8	161.5	8.7	1790	1	USOL_YEAST
9	160	8.6	1433	1	REST_CHICK
10	159.5	8.6	1639	1	LMG1_DROME
11	159.5	8.6	1957	1	YD86_SCHPO
12	157	8.5	483	1	MG_STRPY
13	151.5	8.2	2017	1	MYSN_DROME
14	150	8.1	1938	1	MYSN_ABOIR
15	150	8.1	1966	1	MYSB_CAEEL
16	149	8.0	662	1	TLPA_BACSU
17	148	8.0	2116	1	MYG2_DICDI
18	148	8.0	3110	1	LMA2_HUMAN
19	147.5	8.0	1013	1	SCA4_RICRH
20	147.5	8.0	3210	1	CENF_HUMAN
21	147	7.9	564	1	M12_STRPY
22	147	7.9	1330	1	SMC3_YEAST
23	145	7.8	998	1	SCA4_RICAK
24	143	7.7	1093	1	TMF1_HUMAN
25	142.5	7.7	2869	1	RBP1_PLAVB
26	142	7.7	540	1	CH60_MYCPA
27	142	7.7	3672	1	LMF2_CAEEL
28	141.5	7.6	875	1	TLPI_YEAST
29	141.5	7.6	1786	1	LMB1_MOUSE
30	141	7.6	1427	1	REST_HUMAN
31	140.5	7.6	1011	1	SCA4_RICAF
32	140.5	7.6	1786	1	LMB1_HUMAN
33	140	7.6	1147	1	CGAL_HELPY

34	139.5	7.5	756	1	Y328_MYCCE	049419 mycoplasma
35	139.5	7.5	1935	1	MYSS_CYPCA	090339 cyprinus ca
36	139	7.5	1391	1	MST2_DROHY	008696 drosophila
37	139	7.5	2349	1	TPR_HUMAN	P12270 homo sapien
38	138.5	7.5	944	1	NUFI_YEAST	P32380 saccharomyc
39	138	7.5	724	1	HMMR_HUMAN	075330 homo sapien
40	137.5	7.4	662	1	TLPB_BACSU	P39217 bacillus su
41	137	7.4	697	1	MFPI_LYCES	P93203 lycopersico
42	137	7.4	1500	1	SSP5_STRGN	P16952 streptococc
43	137	7.4	1972	1	MYHB_RABIT	P35748 oryctolagus
44	136.5	7.4	473	1	YVCE_BACSU	P40760 bacillus su
45	136.5	7.4	1189	1	YJH6_YEAST	P47035 saccharomyc

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	210 AA.
OSCL_BORBU	007337			
AC	007337			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BBBI9.			
OG	Borrelia burgdorferi (Lyme disease spirochete).			
OS	Plasmid lp54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID=139;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE-9329332; PubMed-8098841;			
RA	Jauris-Heipke S., Fuchs R., Motz M., Preace-Mursic V., Schwab E.,			
RA	Will G., Wilske B.;			
RT	"Genetic heterogeneity of the genes coding for the outer surface			
RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.;"			
RL	Med. Microbiol. Immunol. 182:37-50(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE-9329332; PubMed-8478108;			
RA	Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,			
RA	Schwab E., Wanner G.;			
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant			
RT	major outer surface protein of Borrelia burgdorferi.;"			
RL	Infect. Immun. 61:2182-2191(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE-94041630; PubMed-8225587;			
RA	Padula S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;			
RT	"Molecular characterization and expression of p23 (OspC) from a North			
RT	American strain of Borrelia burgdorferi.;"			
RL	Infect. Immun. 61:5097-5105(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE-96025162; PubMed-7494039;			
RA	Fukunaga M., Hamase A.;			
RT	"Outer surface protein C gene sequence analysis of Borrelia			
RT	burgdorferi sensu lato isolates from Japan.;"			
RL	J. Clin. Microbiol. 33:2415-2420(1995).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35210 / B31;			
RX	MEDLINE-98065943; PubMed-9403685;			
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,			
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,			
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,			
RA	Peterson J., Kariyave A.R., Quackenbush J., Salzberg S., Hanson M.,			
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,			

Accession	Protein	Length	Similarity
DR EMBL: X03596;	CAA49306.1;	-	EMBL: X03596;
DR EMBL: U01894;	AAI16058.1;	-	EMBL: U01894;
DR EMBL: D49497;	BAA08457.1;	-	EMBL: D49497;
DR EMBL: AE000792;	AAC66329.1;	-	EMBL: AE000792;
DR TIGR: BB819;	-	-	TIGR: BB819;
DR InterPro: IPR001800;	Lipoprotein_6.	-	InterPro: IPR001800;
DR Pfam: PF01441;	Lipoprotein_6; 1.	-	Pfam: PF01441;
DR ProDom: PD001149;	Lipoprotein_6; 1.	-	ProDom: PD001149;
DR ProSite: PS00013;	PROKAR_LIPOPROTEIN; 1.	-	ProSite: PS00013;
KW Outer membrane;	Lipoprotein; Signal; Plasmid; Antigen;	-	Outer membrane;
KW Complete proteome.	-	-	Complete proteome.
FT SIGNAL	1	18	BY SIMILARITY.
FT CHAIN	19	210	OUTER SURFACE PROTEIN C.
FT LIPID	19	19	N-ACYL DILYSERIDE (BY SIMILARITY).
SO SEQUENCE	210 AA;	22340 MM;	7A4FC97BF9177BF CRC64;

QY	1	MACNSGKRGKNTSANSADSVKGPMLTEINKRTIDSNAVLAVAEVALLSTIDETIAKA	60
		::	
Db	17	ICSNNSGKRGKNTSANSADSVKGPMLTEISKRTIDSNAVLAVAEVALLSTIDETIAKA	76
QY	61	IGKRTIHONNGDPTENNHHGSLAGAYALSTLIKOKLDLKGNEGIKERIDAAKCSPTFN	120
		::	
Db	77	IGKRTIHONNGDPTENNHHGSLAGAYALSTLIKOKLDLKGNEGIKERIDAAKCSPTFN	136
		::	
QY	121	KLKEKHTDLGREGVYDADAKFAILKANGRTKGAEBELGKLFESVEVLSRAAEMLANSYK	180
		::	
Db	137	KLKEKHTDLGREGVYDADAKFAILKANGRTKGAEBELGKLFESVEVLSRAAEMLANSYK	196
		::	
QY	181	ELTSPVAESPKRP 194	
		::	
Db	197	ELTSPVAESPKRP 210	

```

RA      Soutschek E. :
RT      "Molecular analysis and expression of a Borrelia burgdorferi gene
RL      encoding a 22 kDa protein (Pc) in Escherichia coli. " ;
RN      Mol. Microbiol. 6:503-509(1992).
[2]
RP      SEQUENCE OF 1-205 FROM N.A.
RC      STRAIN-DK26;
RX      MEDLINE=94075528; PubMed=8253951;
RA      Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K. ;
RT      "Polymorphism in ospC gene of Borrelia burgdorferi and
RT      immunoreactivity of ospC protein: Implications for taxonomy and for
RT      use of ospC protein as a diagnostic antigen. " ;
CC      J. Clin. Microbiol. 31:2570-2576(1993).
CC      - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC      - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC      ANCHOR.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL collaboration -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X62162; CA444093.1; -
DR      EMBL: X73624; CA452003.1; -
DR      InterPro: IPR001800; Lipoprotein_6.
DR      Pfam: PF01441; Lipoprotein_6; 1.
DR      ProDom: PD001149; Lipoprotein_6; 1.
KW      Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT      SIGNAL 1 18
FT      CHAIN 19 212
FT      LIPID 19 19
SQ      SEQUENCE 212 AA; 22499 MW; C206C231FBE2ETD4 CRC64;
                                OUNER SURFACE PROTEIN C.
                                N-ACIL DISULFIDE (BY SIMILARITY).

Query Match      34.1%; Score 631; DB 1; Length 212;
Best Local Similarity 69.9%; Pred. No. 7, 6e-23;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

QY      1 MACNNNGSGDGT -SANSADSVSGPNLTETINKRTTDSNAVLAVKEVALLSIDEIAAK 59
DB      17 ISCNNGSGKGGASSTNPADSDAKGPNLTETISKRTTDSNAFLAVKEVELVLSIDEIAK 76
QY      60 AIGKKIHQNNGLDTENNHNSGLAGVAISTLKOKIDGKN-EGLEKRTIDAAKCSSEF 118
DB      77 AIGKIKINNNGALNANNNGSLAGVAISTLTETLSKKNLEELKRTIDAAKCSSEF 136
QY      119 TNKLEKERTLTGKEGVYDDAKFAILLKANGCTKKGAELEKLFESVEYLSAKAKTEMLANS 178
DB      137 TNKLSGHADLGQDATTDDAKAKAILKTHATYTDKGAKEKDFEVEYGLAKAAQVALNS 196
QY      QY      179 VKELTSPVVAESPKP 194
DB      197 VKELTSPVVAESPKP 212

RESULT      3
ID      VM24.BORHE
AC      P32278;
STANDARD; PRT; 214 AA.

DT      01-OCT-1993 (Rel. 27, Created)
DT      01-OCT-1993 (Rel. 27, Last sequence update)
DE      01-FEB-1994 (Rel. 28, Last annotation update)
DE      VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN      VM24.
OS      Borrelia hermsli.
OC      Plasmid.
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NO      NCBI-TaxID=140;
RN      [1]
RP      SEQUENCE FROM N.A.

```


CC LIGHT CHAIN SUBUNITS (M.C.) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (M.C.-2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
 CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
 CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Y00624; CA66863.1; -
 DR PIR: A27224; A27224.
 DR HSSP: P08799; IMND.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR MYOSIN: Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
 KM Methylation; Alkylation; Phosphorylation; Multigene family.
 FT DOMAIN 1 789
 FT DOMAIN 1 819
 FT DOMAIN 848 1509
 FT DOMAIN 848 1509
 FT DOMAIN 1227 1252
 FT DOMAIN 1253 1509
 FT DOMAIN 1253 1482
 FT DOMAIN 1483 1509
 FT NR_BIND 182 189
 FT DOMAIN 660 682
 FT DOMAIN 766 780
 FT MOD_RES 133 133
 FT MOD_RES 700 700
 FT MOD_RES 1489 1489
 FT MOD_RES 1494 1494
 FT MOD_RES 1499 1499
 FT MOD_RES 1509 1509
 FT MOD_RES 171201 171201
 FT MOD_RES 2CE49BE51173D17E CRC64;
 SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 9.4%; Score 174.5; DB 1; Length 1509;
 Best Local Similarity 24.7%; Pred. No. 0.38;
 Matches 98; Conservative 73; Mismatches 159; Indels 67; Gaps 19;

QY 28 EINKKITSNAVLAVKEVEA-----LSSIDEIAKAKGKIKHQ--NGGLTENNHNHNSL 81
 DB 934 ELQEEFTSANDLLEQKRIKLEAKGELKASLEE--EENRRALDEAKTVKSEENRELODK 990
 QY 82 LAGAVAIITLLKQKLDGKLNKGLKEKIDA--AKKSEFTFNKLEKKTDTGKRGVYDA- 137
 DB 991 YEDEAAHDSLKKEEDLSRE-LRETKDALDAENISITLSKTL--KNTGAGADVKNEL 1047
 QY 138 -DAKPAIKANGTKKGAEELGKLEFSEYVSKAKELANSVKEITSPV----- 186
 DB 1048 DDVATKTIQLEKTKKSLKEELAEQTRAOLEE-EKSGKEAASSKAKOLOQOEDARSEVDSL 1106
 QY 187 -----VASEPKKPSMVNNSG-----KDGNTSANGADESVKG--PNLFEISKIT-----ES 230
 DB 1107 KSKLSAEKSIKTAQDNRDLDLEDERIVYRANVDKOKALEAKTLELDQVATADGOK 1166

QY 231 NAVLAVKEIETLLASIDELATKAIGKIOONGLA--VEAGNGLLAGATTISKLTQ 288
 DB 1167 NA---AAQAOKITLKTQVDE--TK---RRLAEASARLEKERNKALDEVAULTDLDLAE 1218
 QY 289 KLGGLKXSEKLEKTI-----ENARK---CSEDFTKLEGEHAQIGINVTDENAKKAI 338
 DB 1219 RDSGAQQRRLNIRISELOSELENGASSEEYKRLGSELERLEEBLLAOGERRAA- 1277
 QY 339 LITDAKDKGAELKLEFRAVENLAKAKEMANSVK 375
 DB 1278 --AEKNDKRNLELEELRQADDAARDNDKLVKDNKR 1312
 RESULT 6
 ID M5_STRPY STANDARD; PRT; 492 AA.
 AC P02977;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 01-FEB-1991 (rel. 17, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE M PROTEIN, SEROTYPE 5 PRECURSOR.
 GN EMM5 OR SMP5.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88186881; PubMed=3281944;
 RA Miller L., Gray L., Beachey E., Kehoe M.;
 RT "Antigenic variation among group A streptococcal M proteins.
 RT Nucleotide sequence of the serotype 5 M protein gene and its
 RT relationship with genes encoding types 6 and 24 M proteins.";
 RL J. Biol. Chem. 263:5668-5673(1988).
 RN [2]
 RP SEQUENCE OF 43-212 AND 238-250.
 RX MEDLINE=84162039; PubMed=6368549;
 RA Manjula B.N., Acharya A.S., Mische S.M., Fairwell T., Fischetti V.A.;
 RT "The complete amino acid sequence of a biologically active
 RT 197-residue fragment of M protein isolated from type 5 group A
 RT streptococci.";
 RL J. Biol. Chem. 259:3686-3693(1984).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
 CC PHAGOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
 CC -1- SIMILARITY: TO OTHER M PROTEINS.
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: M20374; AAA26976.1; -
 DR PIR: A03501; M5SOMP.
 DR PIR: A28616; A28616.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR003345; M_repeat.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 DR Pfam: PF02370; M_9
 DR PRINTS: PR00015; GPOSANCHOR.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
 DR Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
 KM Transmembrane; Coiled coil; Signal.
 FT SIGNAL 1 42
 FT CHAIN 43 492
 FT DOMAIN 43 466 M PROTEIN, SEROTYPE 5.
 FT DOMAIN 43 466 EXTRACELLULAR (POTENTIAL).


```

FT TRANSMEM 467 486 MEMBRANE ANCHOR.
FT DOMAIN 492 492 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 491 457 GIV/PRO-RICH (CELL WALL-SPANNING).
FT DOMAIN 458 463 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT PROTEINS.
FT DOMAIN 69 103 5 X 7 AA TANDEM REPEATS OF L-K-T-K-N-E-G.
FT REPEAT 69 75 1.
FT REPEAT 76 82 2.
FT REPEAT 83 89 3.
FT REPEAT 90 96 4.
FT REPEAT 97 103 5.
FT CONFLICT 43 43 A -> T (IN REF. 2).
FT CONFLICT 50 50 N -> S (IN REF. 2).
FT CONFLICT 102 102 K -> SNEERKTAELTSEK (IN REF. 2).
FT CONFLICT 208 208 I -> L (IN REF. 2).
SQ SEQUENCE 492 AA: 55085 MW: 873779B6CBD5E27 CRC64;

Query Match 9.3%; Score 172; DB 1; Length 492;
Best Local Similarity 24.9%; Pred. No. 0.14;
Matches 107; Conservative 54; Mismatches 150; Indels 118; Gaps 18;

31 KRTDSNAVLAV-----KEVALLS-----SIDEIAKAIGKIHONNGIDTENN 77
16 KKGTAVALALSLGAGLVVNTNEVSAAYTRGINDPQAKELDKYTELENDHDKTKNEG 75
78 NSGLLAGAVAISTLIKOKLDGLK--NEGK--EKIDAAKCESTETNKKREKHTDLGKE 132
76 -----LKTENEGKLTENEGKLTENEGKLTENEGKLTENEGKLTENEGKLTENEG 121
133 GVT-DADAK-----EALIKANGTK-----KGAELGKLFESVEVLSKAKEMANSVKE 101
122 KELLEREVONTQYNNNTLTKIKNDLTKELNKTROELANKQOESKEKALNELLEKTVE 101
182 LTPVAESPCKPMSVNNNSGKDGNTSANSADSESVKPNLTKITESNAVLAVKEIE 241
182 -----KIKKEBENKETT-----GTAKILDEIVKDKIAKEQENKET-----IG 219
242 TLASIDELATKAIGK--KIQON--GGLAVEAGHNGTLGAVYTSKLITQKLDLKNSEK 298
220 TLKRIIDETVKKLAKESKONIGALKOELAKKDEANKISDSRKGRLDASREAKK 279
299 LKE-----KIENAKKCE-----DPTKLEGHAGLGIENVTDENAKKAII-L 339
280 OLEAEHOKLEBONKISEASRKGRLDASREAKKQLEAQKLEBONKISEASRGLRR 339
340 ITDAADK-----KGAELKLEFKAVENL-----AKRAKE 368
340 DLDASREAKKQVKEALEANSKLALEKLEKLEBESKKTLEKKAELQAKLEAKKAKE 399
369 MLANSVKEI 377
400 QLAQKAEEL 408

RESULT 7
BAG_STRAG STANDARD; PRT; 1164 AA.
AC P27951;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IGA FC RECEPTOR PRECURSOR (BETA ANTIGEN) (B ANTIGEN).
BAG.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxId=1311;
RX MEDLINE=91312121; PubMed=1857207;
RA Jeristrom P.G., Chhatwal G.S., Timmis K.N.;

```

```

RT "The Iga-binding beta antigen of the c protein complex of Group B
RT streptococci: sequence determination of its gene and detection of two
RT binding regions."
RT Mol. Microbiol. 5:843-849(1991).
RN (2)
RP IDENTIFICATION OF IG-LIKE DOMAIN.
RX MEDLINE=97035265; PubMed=8880921;
RA Bateman A., Eddy S.R., Chothia C.;
RL Protein Sci. 5:1939-1942(1996).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59771; CAA42442.1; -.
DR PIR: S15330; RCSOAG.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003599; Ig.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW Cell wall; Transmembrane; Receptor; Repeat; Signal;
KW Immunoglobulin domain.
FT CHAIN 1 37
FT DOMAIN 38 1164 IGA FC RECEPTOR.
FT TRANSMEM 1141 1159 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1160 1164 MEMBRANE ANCHOR (POTENTIAL).
FT DOMAIN 434 534 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 199 438 IG-LIKE DOMAIN.
FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).
FT DOMAIN 827 945 IGA-BINDING (POTENTIAL).
FT DOMAIN 1131 1137 PRO-RICH REPEATS.
FT CONSERVED IN GRAM-POSITIVE COCCI SURFACE
SQ SEQUENCE 1164 AA: 131051 MW: 650DE94AF720A5474 CRC64;

Query Match 8.9%; Score 164.5; DB 1; Length 1164;
Best Local Similarity 25.1%; Pred. No. 0.81;
Matches 95; Conservative 55; Mismatches 147; Indels 81; Gaps 18;

QY 6 SGRDGTNSANSADSESVKPNLTKITESNAVLAVK---EVEALLSIDEIAKAIG 62
DB 128 SHQKNEFKTKIDETNSDALLEENQFNETNR-LHIKQHEVEKDKAKQKQTLKQSD 186
QY 63 KTIHONNGLTENNHNHNSLAGAVAISTLIKOKLDGLKNGKAKKIDAAKCESTETNKL 122
DB 187 TKVDLSN-IDKELNHOKS-----QVEKMAEQK--GITNED-----KDSM 222
QY 123 KEKHTDLGREGVVDADAEKAILKANGTKTGAEELGKLFESVEVLSKAKEMANSVKEI 182
DB 223 LKRIEDIRKQ-AQADAKKE-----DAEVKVRRELGLTFSS-----TKAG-----L 261
QY 183 TSPVAESPCKPMSVNNNSGKDGNTSANSADSESVKPNLTKITESNAVLV----- 235
DB 262 DQEIQEHVKKETSSEETQKVDEHYANSI-QNLAKQSLSELDKATNTQEQATQVKNQFLEN 320
QY 236 --AVKEIETLASIDELATKAIGKIKIQONGGLAVEAHNG-----TLAAGVYTSKLITQ 288
DB 321 AQRKLEIQLPIKETTNYLVYRAMSESLQ---VEKELHNSEANLEDIIVAASKELVREYEG 377
QY 289 KLDGLKNSKIKREKIEVA-----KKCSDEFTKLEGEHAQGLIENVTDENAKKAIIITDAA 344
DB 378 KLDQSKNLPKLQLEEAHNSKLVQVYEDFRKKRTS-----EGVYTKKKVKKRDL---AA 428

```

QY 345 KDKGAELERKFAVENL 362
Db 429 NENNOKIE-LTSPENI 445

RESULT 8
USOL_YEAST
ID USOL_YEAST STANDARD: PRT: 1790 AA.
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.
GN USOL OR INT1 OR YDL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-X2180-1A;
RX MEDLINE-91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RT J. Cell Biol. 113:245-260(1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54378; CAA38253.1;
CC EMBL; L03188; AAB00143.1;
CC EMBL; 053668; AAB66559.1;
CC PIR; A38455; A38455.
CC HSSP; P80220; IDIP.
CC SGD; S0002216; USOL.
DR InterPro: IPR002017; Spectrin.
DR Transports; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT DOMAIN 1 724 GLOBULAR HEAD.
FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
FT CONFLICT 847 847 G -> E (IN REF. 2).
FT CONFLICT 924 924 E -> K (IN REF. 2).
FT CONFLICT 1253 1253 V -> I (IN REF. 2).
FT CONFLICT 1319 1319 I -> V (IN REF. 2).
FT CONFLICT 1461 1461 N -> S (IN REF. 2).

FT CONFLICT 1581 1581 G -> S (IN REF. 2).
FT CONFLICT 1600 1600 I -> V (IN REF. 2).
FT CONFLICT 1661 1661 R -> S (IN REF. 2).
FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 8.7%; Score 161.5; DB 1; Length 1790;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 99; Conservative 72; Mismatches 178; Indels 107; Gaps 16;

QY 3 CNGSGKDG-----NTSANSADSEYKGGNLEIKKITTSDNANVLAAYEVALLS 52
Db 929 CNNLSKEKEHISKELEVEYKSRFOSHNLV--AKTEKRLKSILANNYKMQA--ENESLKA 984
QY 53 IDEIAKAIKGIKTHONNGID--TENNNGSLAGAY-----ASTL-----TKOR 95
Db 985 VEESKNESSTIQLSNLONKIDSMSQEKENFOIERGSIENKIBOLKTTISDEQTKREILTSK 1044
QY 96 LDGLKNEG-----LKERIDAKKCSFTFKLAE-----KHTDLGKEGVTD 136
Db 1045 SSSSDEYSESQISLKEKLETTATTANDENVKRISELTRETEAEALAAVKNLKNLETRK 1104
QY 137 ADAPKAILK-----ANGTKTK-----GAEELGKLESEYEVLSKAA 171
Db 1105 LETSEKALKEVKNENDEHLKEEKIQLEKREAVETKQOINSLANESLEKEHEDLAOLKRY 1164
QY 172 KEMLANSVKELTSPV--VASPKKPSMVNNSGKDQNTSANSADSEYKGGNLEISKITE 229
Db 1165 EQGLANERQVNEISQINDETITSTOENESIKRKDELEGEVKAARSTSEQSNLKSE 1224
QY 230 SNAVVLAVKEIE-----TLASIDELATKAIGKIIQONGLAAYEAHNTLLAGAT 281
Db 1225 IDALNTQIKELKKNETNEASILEISKVSSEYV--KINE--LDQCNFK----- 1270
QY 282 ISKLTITKLGKLNSEKLEKINAKKCEDPTFKKIEGSHAOLGIENVTENAKKAILIT 341
Db 1271 -EKVESLEKELKASEDKSKSYTELQEKSEKIEKELDATTETLKIQLKRTWLSK----- 1324
QY 342 DAAKDKGAELERKFAVENLAKAAKEMLANSVKEL 377
Db 1325 --AKESSESLRLKTSSEERKNAEEQLEKLNEL 1358
RESULT 9
REST CHICK
ID REST CHICK STANDARD: PRT: 1433 AA.
AC 042184; 042228; 057563; 057564;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).
GN RSN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98137792; PubMed=9469933;
RA Griparic L., Volosky J.M., Keller T.C. III;
RT "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RT Gene 206:195-208(1998).
RN [2]
RP SEQUENCE OF 17-1119 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC TISSUE-Pectoralis muscle.
RA Griparic L., Keller T.C. III;
RT "Identification and expression of two novel CLIP-170/Restin isoforms
RT expressed predominantly in muscle.";
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
 CC CYTOSKELETON (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL: AF014012; AAC60344.1; -;
 DR EMBL: AF020764; AAC60345.1; -;
 DR EMBL: AF045650; AAC03547.1; -;
 DR EMBL: AF045651; AAC03548.1; -;
 DR InterPro: IPR000938; CAP-GLY;
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF01302; CAP_GLY; 2.
 DR SMART: SM00343; Znf_C2HC; 1.
 DR PROSITE: PS00845; CAP_GLY_1; 2.
 KM Cytokeleton: Microtubules; Colled coil; Alternative splicing.
 FT DOMAIN 79 121
 FT 144 207 SER-RICH.
 FT 207 277 CAP-GLY 2.
 FT 335 332 SER-RICH.
 FT 351 1353 COILED COIL (POTENTIAL).
 FT DOMAIN 1414 1427 CCHC-BOX.
 FT VARSPLIC 458 492 MISSING (IN SHORT ISOFORM).
 FT VARSPLIC 458 492 TOTKELHARKELOSLEKTRADKQRELETR ->
 FT VARSPLIC 803 803 RKRQISEDPEN (IN ISOFORM CLIP-170(11)).
 FT VARSPLIC 458 458 T -> GGSSKVS (IN ISOFORM CLIP-170(11)).
 FT VARSPLIC 309 309 S -> RKRQISEDPEN (IN ISOFORM CLIP-
 FT CONFLICT 309 309 170(11+35)).
 FT CONFLICT 440 440 K -> R (IN REF. 2; AAC03547).
 FT CONFLICT 440 440 E -> V (IN REF. 2; AAC03548).
 SQ SEQUENCE 1433 AA; 161026 MW; 5631CE8683459E23 CRC64;

Query Match 8.6%; Score 160; DB 1; Length 1433;
 Best Local Similarity 25.9%; Pred. No. 1.6; Indels 58; Gaps 17;
 Matches 96; Conservative 58; Mismatches 158; Indels 58; Gaps 17;

QY 28 EINKKTTSDNAVLAVKEVEALLSSIDEIAKAKIGKRIHONNGIDTENNHGSLLAGAYA 87
 DB 524 ESSKHIDVDTSLLOEI-----SLQEKMA-AAGKE-HQREMSLKEKFESESEALRKE 577
 DB 88 ISLT-IKQKIDGLKNELEKRIKIDAKR-----CSEFTFNKIKE---KRTDLGKE----- 132
 QY 578 IKTLASNEKMGKENELEKTKRLHANKENSVDTELMKSKLESATASHQQAWEELKVSFNK 637
 QY 133 GVTDAADAKKAILKANGTKTGAEELGKLFESVEVLAKAKEMLANSVKELTSPVAVESPK 192
 DB 638 GVGAGTAEFELKTMQKVLDE--NEMSNIKIKOENESQHLKELEALKAKILEETEE 695
 QY 193 KPSMVNNSGKDGTNSANSADSVKGNPLTEISKITESNAVLAVKEIEFTLLASIDELAT 252
 DB 696 KEOTLEL-----LKAKLESVEDQHLVEMEDTLNQLQAEIKVELDVLQAKCNE-QT 746
 QY 253 KAIGKTIQONGGLAVENGHGTL-----LAGAYTSKILITQKL-DGLKNSERLAKIKENAK 307
 DB 747 KILGSLTQO-----IRSEKSLDLALQKANSSEKLEIKLSQQLAAEQIONLETER 801
 QY 308 KCEDEFTKLEGEHAQGIENVTDENAKKAILITDAKDKGAEL-----EKLFKAV---E 360
 DB 802 --VSNLTKELOQKQK-----LDLEKNLSAVNOVKDSLKEKLOLLKEKFTSAVNGAE 852
 QY 361 NLAKAKEML 370
 DB 853 NAQRAAQETI 862

RESULT 10
 ID LMGI_DROME STANDARD; PRT; 1639 AA.
 AC P15215; 024373; Q9VT18;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 20-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE LAMININ GAMMA-1 CHAIN PRECURSOR (LAMININ B2 CHAIN).
 GN LANB2 OR LAMC1 OR LAMG1 OR CG3322.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7227;
 RX SEQUENCE FROM N.A.
 RC STRAIN=CANTON-S, AND OREGON-R;
 RX MEDLINE=91299161; PubMed=1840513;
 RA Chl H.-C., Jumiaga D., Wang S.Y., Hui C.-F.;
 RT "Structure of the Drosophila gene for the laminin B2 chain.";
 RL DNA Cell Biol. 10:451-466(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;
 RX MEDLINE=89109164; PubMed=2912972;
 RA Chl H.-C., Hui C.-F.;
 RT "Primary structure of the Drosophila laminin B2 chain and comparison
 RT with human, mouse, and Drosophila laminin B1 and B2 chains.";
 RL J. Biol. Chem. 264:1543-1550(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90037237; PubMed=2808533;
 RA Montell D.J., Goodman C.S.;
 RT "Drosophila laminin: sequence of B2 subunit and expression of all
 RT three subunits during embryogenesis.";
 RL J. Cell Biol. 109:2441-2453(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amenlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
 RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glisler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen R.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weissbach J.,

FT	DOMAIN	710	1049	6.5 X LAMININ EGF-LIKE REPEATS (DOMAIN II).
FT	DOMAIN	710	743	LAMININ EGF-LIKE 5 (C-TERMINAL).
FT	DOMAIN	744	792	LAMININ EGF-LIKE 6.
FT	DOMAIN	793	846	LAMININ EGF-LIKE 7.
FT	DOMAIN	847	901	LAMININ EGF-LIKE 8.
FT	DOMAIN	902	955	LAMININ EGF-LIKE 9.
FT	DOMAIN	956	1003	LAMININ EGF-LIKE 10.
FT	DOMAIN	1004	1049	LAMININ EGF-LIKE 11.
FT	DOMAIN	1050	1609	DOMAIN II AND I.
FT	DOMAIN	1087	1109	COILED COIL (POTENTIAL).
FT	DOMAIN	1144	1247	COILED COIL (POTENTIAL).
FT	DOMAIN	1306	1627	COILED COIL (POTENTIAL).
FT	DISULFID	299	308	BY SIMILARITY.
FT	DISULFID	301	322	BY SIMILARITY.
FT	DISULFID	324	333	BY SIMILARITY.
FT	DISULFID	336	356	BY SIMILARITY.
FT	DISULFID	359	368	BY SIMILARITY.
FT	DISULFID	361	384	BY SIMILARITY.
FT	DISULFID	387	396	BY SIMILARITY.
FT	DISULFID	399	411	BY SIMILARITY.
FT	DISULFID	414	426	BY SIMILARITY.
FT	DISULFID	416	432	BY SIMILARITY.
FT	DISULFID	434	443	BY SIMILARITY.
FT	DISULFID	446	458	BY SIMILARITY.
FT	DISULFID	461	475	BY SIMILARITY.
FT	DISULFID	463	482	BY SIMILARITY.
FT	DISULFID	484	493	BY SIMILARITY.
FT	DISULFID	496	511	BY SIMILARITY.
FT	DISULFID	511	553	BY SIMILARITY.
FT	DISULFID	744	753	BY SIMILARITY.
FT	DISULFID	746	760	BY SIMILARITY.
FT	DISULFID	762	771	BY SIMILARITY.
FT	DISULFID	774	790	BY SIMILARITY.
FT	DISULFID	793	801	BY SIMILARITY.
FT	DISULFID	795	811	BY SIMILARITY.
FT	DISULFID	814	823	BY SIMILARITY.
FT	DISULFID	826	844	BY SIMILARITY.
FT	DISULFID	847	861	BY SIMILARITY.
FT	DISULFID	849	868	BY SIMILARITY.
FT	DISULFID	871	880	BY SIMILARITY.
FT	DISULFID	883	899	BY SIMILARITY.
FT	DISULFID	902	919	BY SIMILARITY.
FT	DISULFID	928	937	BY SIMILARITY.
FT	DISULFID	940	953	BY SIMILARITY.
FT	DISULFID	956	968	BY SIMILARITY.
FT	DISULFID	958	975	BY SIMILARITY.
FT	DISULFID	977	986	BY SIMILARITY.
FT	DISULFID	989	1001	BY SIMILARITY.
FT	DISULFID	1004	1016	BY SIMILARITY.
FT	DISULFID	1006	1022	BY SIMILARITY.
FT	DISULFID	1024	1033	BY SIMILARITY.
FT	DISULFID	1036	1047	BY SIMILARITY.
FT	DISULFID	1050	1050	INTERCHAIN (PROBABLE).
FT	DISULFID	1053	1053	INTERCHAIN (PROBABLE).
FT	DISULFID	1631	1631	INTERCHAIN (PROBABLE).
FT	CARBOHYD	147	147	N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match Best Local Similarity 8.6%; Score 159.5; DB 1; Length 1639; Matches 105; Conservative 58; Mismatches 162; Indels 115; Gaps 19.

2 ACNNSGKDGWTSNNSAD-----ESVKGNGNLEIRKTKTSDSNAVLAWEALLSSIDE 55
 1107 ARDNGSGGGTYAEVIDDLKHLDSYR-EHLVASAKFOADNGEIDRARONYTILQDITE 1165

56 IAAKAIKTKTHQNGCLDTEENNNGSLAGAYASTLLKQIKDKNGKNGLEKIDDAKKS 115
 1166 NAKKEL-----QQAALDLINDEGAQALRAKESVGEQD-----SEQISDISREA 1210

116 EFTFNKLK-EKTDLDKKEGVTDADKKKEILAN-----GYTKG--AEELGK 159
 1211 RAADLTFEEAFQDFKND-----YTKKKK----- 11

QY 316 KLEGEHAOLGIENVTDENAKKAILITD-----AAKDKGALEK-LFRAVENTL- 362
 Db 1587 ELKAOHEELEDQLEEDAKLELEVMKQALRSQFERDILAKEGAEKRGVLKQKDLDE 1646
 QY 363 -----AKAKEMLANSVKEL 377
 Db 1647 TELDEERKORTAIVASKKKEGLKEI 1673
 RESULT 14
 MYS_AEOIR STANDARD: PRT: 1938 AA.
 AC P24733;
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MYOSIN HEAVY CHAIN, STRIATED MUSCLE.
 OS Aequipecten irradians (Bay scallop).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
 OC Pectinoidea; Pectinidae; Argopecten.
 NC NCBI_TaxID=31199;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adductor muscle;
 RX MEDLINE=92011595; PubMed=1917970;
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
 RT Complete primary structure of a scallop striated muscle myosin heavy
 RT chain. Sequence comparison with other heavy chains reveals regions
 RT that might be critical for regulation.;
 RL J. Biol. Chem. 266:18469-18476(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adductor muscle;
 RX MEDLINE=91088319; PubMed=2263488;
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
 RT Nucleotide sequence of full length cDNA for a scallop striated
 RT muscle myosin heavy chain.;
 RL Nucleic Acids Res. 18:7158-7158(1990).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
 RX MEDLINE=94173332; PubMed=8127365;
 RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalborkis V.N.,
 RA Szent-Gyorgyi A.G., Cohen C.;
 RT Structure of the regulatory domain of scallop myosin at 2.8-A
 RT resolution.;
 RL Nature 368:306-312(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
 RX MEDLINE=96419133; PubMed=8805510;
 RA Houdusse A., Cohen C.;
 RT Structure of the regulatory domain of scallop myosin at 2-A
 RT resolution: Implications for regulation.;
 RL Structure 4:21-32(1996).
 RN -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL, X55714; CAA39247.1; -

DR PIR: S13557; S13557.
 DR PIR: A40997; A40997.
 DR PDB: 1SCM; 30-APR-94.
 DR PDB: 1WDC; 11-JUL-96.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR002017; Spectrin.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00663; myosin_head; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS5096; IQ; 1.
 DR KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
 FT DOMAIN 1 777
 FT DOMAIN 778 805
 FT DOMAIN 836 1938
 FT DOMAIN 836 1938
 FT NP_BIND 176 183
 FT MOD_RES 693 693
 FT MOD_RES 703 703
 FT HELIX 778 821
 FT TURN 822 823
 FT TURN 825 833
 FT HELIX 834 835
 SO SEQUENCE 1938 AA; 22821 MW; A5CCEA127D1A4896 CRC64;

Query Match 8.1%; Score 150; DB 1; Length 1938;
 Best Local Similarity 20.4%; Pred. No. 6.4;

Matches 81; Conservative 65; Mismatches 166; Indels 86; Gaps 12;

QY 18 DESKGNLREINKKINDSNVLLAVEKEALLSSIDEIAKATGKRIHONGHDT---- 73
 Db 925 DEEDDAADLEGIRKKMKADNA-----NLKKIDGDLLENTLOKAQDRAHKNOISTLOGE 978
 QY 74 ---ENNHSILGAVAIISTLIKOKLDLKNGLKRIKIDAKKSETFTNKKKHTDLG 130
 Db 979 ISQDDEHIGKLNKKKRLLEANKKTSLSQAE--EDCNHNLKILAKLEQALDELENLE 1036
 QY 131 KEGYTDAKAEALIKANGTKTGAEEIGKLFESYEVLSKAKEMLANSVKELTSPVAES 190
 Db 1037 REKKVRGDVEKA-----RKRVQDLKSTQENVEDLERVARELEN-----VRRK 1080
 QY 191 PKKPSMNNNGSKGNTSANSADSEVK--GRPLTETISKRI--TESNAVYLVAVEIETLLASI 247
 Db 1081 EAELISLNSKLEDEQNLVLSQLOKRIKELQARIELELELEERNAKRAKVEQRAELREL 1140
 QY 248 DELATKAIIGKKIOONG-----LAVENGHTLLAGAYTISKL----- 285
 Db 1141 EEL-----GERLDEAGATSAQIELNKKREAEELKIRDELEASLQHEAQISALRKKHQD 1195
 QY 286 -ITOKLDGLNKKSEKLEKRIENAKK-----CSDFPKKLEGEHAQ 323
 Db 1196 AANEMADQVQDLQVKKLEKDKDLKREMDLESQTHNNKKNKGCSEKVKQFESQKMSD 1255
 QY 324 LGIENVTDENAKKAILITDAKDKGALEKLFRAVEN 361
 Db 1256 L--NARLEDSORSINELQSOKSRLQAENSDLTROLED 1290
 RESULT 15
 MYS_CAEEL STANDARD: PRT: 1966 AA.
 ID MYS_CAEEL
 AC P02566;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MYOSIN HEAVY CHAIN B (MHC B).

GN UNC-54 OR MYO-4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RX SEQUENCE FROM N.A.
 RA MEDLINE=83273600; PubMed=6576334;
 RT "Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 RT myosin heavy chain gene are not separated by introns.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RP SEQUENCE OF 850-1966 FROM N.A.
 RX MEDLINE=82272395; PubMed=7202124;
 RA McLaughlin A.D., Karn J.;
 RT "Periodic charge distributions in the myosin rod amino acid sequence
 RT match cross-bridge spacings in muscle.";
 RL Nature 299:226-231(1982).
 [3]
 RP SEQUENCE OF 1876-1966 FROM N.A.
 RX MEDLINE=83232892; PubMed=6571695;
 RA Wallis N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
 RA Waterston R.H.;
 RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber
 RT nonsense mutations via altered transfer RNA.";
 RL Cell 33:575-583(1983).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS
 CC P.TM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
 CC ELEGANS.
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 DR EMBL: J01050; AAA28124.1; -
 DR EMBL: V01494; CAA24738.1; -
 DR PIR: A02992; MWK.
 DR HSSP: P08799; IMND.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head.1.
 DR Pfam: PF01576; Myosin_tail.1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head.1.
 DR SMART: SM00242; MYSC.1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
 FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).
 FT DOMAIN 1165 1176 HINGE.
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMM).

FT NE_BIND 177 184 ATP (BY SIMILARITY).
 FT DOMAIN 665 687 ACTIN-BINDING.
 FT DOMAIN 769 783 ACTIN-BINDING.
 FT MOD_RES 128 128 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 705 705 ALKYLATION (SH-1).
 FT MOD_RES 715 715 ALKYLATION (SH-2).
 FT CONFLICT 1337 1337 E -> R (IN REF. 2).
 FT CONFLICT 1880 1880 I -> L (IN REF. 2).
 SQ SEQUENCE 1966 AA; 225125 MW; B6F0BB2FE27B67F CRC64;

Query Match 8.1%; Score 150; DB 1; Length 1966;
 Best Local Similarity 22.7%; Pred. No. 6.5;
 Matches 90; Conservative 76; Mismatches 149; Indels 82; Gaps 20;

QY 16 SADESVMKGNLEIMKTKITDSNAVLAVK---EVALLSIDET--AKAIGKTIHQ--NN 69
 DB 1358 SLEELIEGKN--ELIRQSKANADIQOMKARPEGGLKA--DELDARKROAKINELQ 1414
 QY 70 GLDTENNNGSL-LAGAVAISTLIKOKLGLKNEGLKEKIDAKKCEFTNKLKEKHTD 128
 DB 1415 ALDANKSKNASLEKTKSRVLGDDAQYDVERANGVASALEKKKGPKKIIDEMRKTTD 1474
 QY 129 LKEGVTDADAKAELTKANGT---KTGADELGLFESEVYLSKAKEMLANSVKELTSP 185
 DB 1475 LAAE---LDGAQRDLKNTSTDLFEAKNAQE--ELAEVVEGLIRENKS-LSQETIKDLTDQ 1527
 QY 186 VVAESPKRPSMVNNSGKDNTSANSADSVKGNL--TEISKITYESNAVLAVKEIETL 243
 DB 1528 L-----GEGRSVHEMOKIIRLEIEKELEQHALDENVA-ALAESEKVL 1571
 QY 244 LASIDELATKA-IGKKIO-----NGLAVBAGHNGTLGAVTIS 283
 DB 1572 RAQVVSQIRSEIEKRIQKEEKEEFTNKNHARALESMQSLFTEAKGAKELLR----- 1625
 QY 284 KLITQKLDGLKNEKTKIEKINAKKSEDPFKLGEHAQGINVDENKKAIIITDA 343
 DB 1626 --IKKLEGG--DINLELALDHAKKANADAKKNLR-----YQGVRELDQVEE 1671
 QY 344 AKDKGALELEKLFKAVE--NLAKAKE--MLANSVKE 376
 DB 1672 EORNGADTRQOFNNAEKRAVLQSEKEBELLVANAEAE 1708

Search completed: March 18, 2002, 10:11:49
 Job time: 967 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:53 ; Search time 124.19 Seconds
(without alignments)
444.035 Million cell updates/sec

Title: US-09-596-746a-30

Perfect score: 1850
Sequence: 1 MACNSGKDGNTSANSADSE.....AVENLAKAKKELANSVKEL 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17:*
1: SP_Archea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Hit No.	Score	Query Match	Length	DB ID	Description
1	943	51.0	192	2	Q9S3P3
2	927	50.1	191	2	P70818
3	921	49.8	192	2	Q9R7B1
4	908	49.1	193	2	Q9R8S3
5	893	48.3	200	2	Q9R0R9
6	885	47.8	192	2	Q9R7B2
7	876	47.4	193	2	P94234
8	876	47.4	194	2	Q31122
9	876	47.4	201	2	Q9R0R8
10	863	46.6	191	2	Q44726
11	853	46.1	177	2	Q44999
12	843.5	45.6	177	2	Q9REH7
13	841	45.5	195	2	Q31123
14	829	44.8	175	2	Q9R7B4
15	826	44.6	178	2	Q44995
16	823	44.5	178	2	Q44979
17	802	43.4	163	2	Q9R7A9
18	774.5	41.9	211	2	Q44720
19	773.5	41.8	193	2	P94237

20	773	41.8	159	2	Q9R7B0	Q9R7B0 borrelia bu
21	739	39.9	178	2	Q44994	Q44994 borrelia bu
22	709	38.3	210	2	Q44719	Q44719 borrelia bu
23	707	38.2	192	2	Q9S3P2	Q9S3P2 borrelia bu
24	696.5	37.6	211	2	Q44977	Q44977 borrelia bu
25	692.5	37.4	211	2	Q49576	Q49576 borrelia af
26	690.5	37.3	193	2	P94242	P94242 borrelia bu
27	686.5	37.1	211	2	Q44976	Q44976 borrelia bu
28	686	37.1	190	2	P94244	P94244 borrelia bu
29	686	37.1	212	2	Q9KIM5	Q9KIM5 borrelia af
30	685.5	37.1	209	2	Q44671	Q44671 borrelia af
31	683.5	36.9	191	2	P94245	P94245 borrelia bu
32	683	36.9	193	2	Q31115	Q31115 borrelia bu
33	681	36.8	212	2	Q44705	Q44705 borrelia bu
34	675	36.5	194	2	P94247	P94247 borrelia bu
35	670	36.2	212	2	Q44727	Q44727 borrelia bu
36	667	36.1	201	2	P96571	P96571 borrelia ja
37	666.5	36.0	193	2	P94231	P94231 borrelia bu
38	666.5	36.0	211	2	Q926C7	Q926C7 borrelia va
39	666	36.0	201	2	P96573	P96573 borrelia ja
40	664.5	35.9	193	2	P94233	P94233 borrelia bu
41	664.5	35.9	202	2	P96508	P96508 borrelia af
42	664.5	35.9	209	2	Q44883	Q44883 borrelia bu
43	664.5	35.9	211	2	Q49577	Q49577 borrelia af
44	664	35.9	203	2	O50624	O50624 borrelia af
45	664	35.9	212	2	Q44670	Q44670 borrelia af

ALIGNMENTS

RESULT	ID	Q9S3P3	PRELIMINARY:	PRT:	192 AA.
Q9S3P3	Q9S3P3	Q9S3P3			
AC	Q9S3P3	Q9S3P3			
DT	01-MAY-2000	(TREMblrel. 13, Created)			
DT	01-MAY-2000	(TREMblrel. 13, Last sequence update)			
DT	01-JUN-2001	(TREMblrel. 17, Last annotation update)			
DE	OUTER SURFACE PROTEIN C (FRAGMENT).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID-139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-IP2;				
RX	MEDLINE-96296448; PubMed-8709845;				
RA	Lavey I., Gibbs C.P., Schuster R., Dorner F.;				
RT	"Evidence for lateral transfer and recombination in OspC variation in Lyme disease Borrelia."				
RL	MOL. Microbiol. 18:257-269(1995).				
DR	EMBL; L42887; AAB36995.1; "				
DR	InterPro; IPR001800; Lipoprotein_6.				
DR	Pfam; PF01441; Lipoprotein_6; 1.				
DR	ProDom; PD001149; Lipoprotein_6; 1.				
FT	NON_TER				
FT	1				
FT	192				
FT	192				
SO	SEQUENCE	192 AA; 20287 MW; 11846F7AC84C7E3D CRC64;			
Query Match	Best Local Similarity	51.0%; Score 943; DB 2; Length 192;			
Matches	190; Conservative	1; Mismatches	1; Indels	0; Gaps	0;
Qy	3	CNSGKDGNTSANSADSEVKGNLTETIKKTTSDNAVLAVKEVALLSSIDEIAKAIG 62			
Db	1	CNSGKDGNTSANSADSEVKGNLTETIKKTTSDNAVLAVKEVALLSSIDEIAKAIG 60			
Qy	63	KKHONNGLDTEENNNGSLAGAVAIPTLIRKOKLDGLKNEGLKEKIDAKKCSFTNKL 122			
Db	61	KKHONNGLDTEENNNGSLAGAVAIPTLIRKOKLDGLKNEGLKEKIDAKKCSFTNKL 120.			
Qy	123	KEKHTDLGKEGVTDDADAEALIKANGTKTKGAEEUGKLFESVEVLSKAKKELANSVKEL 182			

```

Db 121 KEKHTDLGEGVTDADAKKALIKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180
QY 183 TSPVAESPKKP 194
Db 181 TSPVAESPKKP 192

RESULT 2
P70818
ID P70818 PRELIMINARY: PRT: 191 AA.
AC P70818:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN (FRAGMENT).
OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2-1498 CA4;
RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: I81131; AAB06569.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
SQ SEQUENCE 191 AA; 20126 MW; D2B9B1C82B4DC3C0 CRC64;

Query Match 50.1%; Score 927; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 1.4e-36;
Matches 186; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 NNSGKDGNTSANSADSVKGNLTETIKTTDSNAVLAVKEVEALLSSIDEIAAK 63
ID [1]
AC Q9R7B1 PRELIMINARY: PRT: 192 AA.
QY 64 KIHQNNGLDTENNHNSLAGAVAI STLKOKLDGKNEGKFKEDIAAKKCSFTFNK 123
Db 61 KIHQNNGLDTENNHNSLAGAVAI STLKOKLDGKNEGKFKEDIAAKKCSFTFNK 120
QY 124 EKHTDGLGEGVTDADAKKALIKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 183
Db 121 EKHTDGLGEGVTDADAKKALIKNTGKTGAEELGKLFESVYLSKAKEMLANSVKEL 180
QY 184 SPVAESPKKP 194
Db 181 SPVAESPKKP 191

RESULT 3
Q9R7B1
ID Q9R7B1 PRELIMINARY: PRT: 192 AA.
AC Q9R7B1:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RA MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RA "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91798; AAB81895.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 20297 MW; 6770502A20AA764 CRC64;

Query Match 49.8%; Score 921; DB 2; Length 192;
Best Local Similarity 97.9%; Pred. No. 2.7e-36;
Matches 186; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETIKTTDSNAVLAVKEVEALLSSIDEIAAK 60
Db 3 ISCNNSGKDGNTSANSADSVKGNLTETISKRTTDSNAVLAVKEVEALLSSIDEIAAK 62
QY 61 IGKTIHQNGLDTENNHNSLAGAVAI STLKOKLDGKNEGKFKEDIAAKKCSFTFN 120
Db 63 IGKTIHQNGLDTENNHNSLAGAVAI STLKOKLDGKNEGKFKEDIAAKKCSFTFN 122
QY 121 KKEKHTDGLGEGVTDADAKKALIKNTGKTGAEELGKLFESVYLSKAKEMLANSVK 180
Db 123 KKEKHTDGLGEGVTDADAKKALIKNTGKTGAEELGKLFESVYLSKAKEMLANSVK 182
QY 181 ELTSPVAES 190
Db 183 ELTSPVAES 192

RESULT 4
Q9R53
ID Q9R53 PRELIMINARY: PRT: 193 AA.
AC Q9R53:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OC1;
RA MEDLINE=99091544; PubMed=9872945;
RA Wang I.N., Dykhuizen D.E., Oju W., Dunn J.J., Bosler E.M., Luft B.J.;
RT "Genetic diversity of ospC in a local population of Borrelia
RL burgdorferi sensu stricto.";
RL Genetics 151:15-30(1999).
DR EMBL: AF029860; AAB86543.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 20502 MW; 5EFD5A5F8986D1E CRC64;

Query Match 49.1%; Score 908; DB 2; Length 193;
Best Local Similarity 97.9%; Pred. No. 1.1e-35;
Matches 183; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETIKTTDSNAVLAVKEVEALLSSIDEIAAK 60
Db 7 ISCNNSGKDGNTSANSADSVKGNLTETISKRTTDSNAVLAVKEVEALLSSIDEIAAK 66
QY 61 IGKTIHQNGLDTENNHNSLAGAVAI STLKOKLDGKNEGKFKEDIAAKKCSFTFN 120

```

DB 67 IGRKIHONNGLDTEENNHNHNSLLAGAVAIISYLKOKLDGLKNEGLKEKIDAKKCEFTFN 126
QY 121 KLEKHTDGLKEGVTDADAKAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 127 KLEKHTDGLKEGVTDADAKAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 186
QY 181 ELTSPV 187
DB 187 ELTSPV 193

RESULT 5
ID Q9R09 PRELIMINARY: PRT: 200 AA.
AC Q9R09: 01-MAY-2000 (TREMUREL. 13, Created)
DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)
DE 01-JUN-2001 (TREMUREL. 17, Last annotation update)
OS OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
GN OSCP.
RA Borrelia burgdorferi (Lyme disease spirochete).
PL plasmid cp26.
OX NCBI_TaxID=139:
RN Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
RP SEQUENCE FROM N.A.
RC STRAIN=27;
RX MEDLINE=20002545; PubMed=10531219;
RT Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT "Population dynamics of a naturally occurring heterogeneous mixture of
RT Borrelia burgdorferi clones.";
RL Infect. Immun. 67:5709-5716(1999).
DR EMBL: AF074464; AAD23911.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KW Plasmid.
FT NON_TER 200
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BB6B6B CRC64;

Query Match 48.3%; Score 893; DB 2; Length 200;
Best Local Similarity 97.8%; Pred. No. 5.7e-35;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNSGKDGNTSANSADSVKGNLLEINKKITDTSNAVLAVKEVALLSSIDEIATAKA 60
DB 17 ISCNNSGKDGNTSANSADSVKGNLLEINKKITDTSNAVLAVKEVALLSSIDEIATAKA 76
DB 61 IGRKIHONNGLDTEENNHNHNSLLAGAVAIISYLKOKLDGLKNEGLKEKIDAKKCEFTFN 120
DB 77 IGRKIHONNGLDTEENNHNHNSLLAGAVAIISYLKOKLDGLKNEGLKEKIDAKKCEFTFN 136
QY 121 KLEKHTDGLKEGVTDADAKAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 180
DB 137 KLEKHTDGLKEGVTDADAKAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVK 196
QY 181 ELTS 184
DB 197 ELTS 200

RESULT 6
ID Q9R7B2 PRELIMINARY: PRT: 182 AA.
AC Q9R7B2: 01-MAY-2000 (TREMUREL. 13, Created)
DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)
DE 01-JUN-2001 (TREMUREL. 17, Last annotation update)
OS OUTER SURFACE PROTEIN C (FRAGMENT).
GN Borrelia burgdorferi (Lyme disease spirochete).
RA Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139:
RN Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
RP SEQUENCE FROM N.A.
RC STRAIN=27;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=27;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91797; AAB81894.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 182
SQ SEQUENCE 182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match 47.8%; Score 885; DB 2; Length 182;
Best Local Similarity 98.9%; Pred. No. 1.2e-34;
Matches 180; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 GKDNTSANSADSVKGNLLEINKKITDTSNAVLAVKEVALLSSIDEIATAKAIGRKIH 66
DB 1 GKDNTSANSADSVKGNLLEINKKITDTSNAVLAVKEVALLSSIDEIATAKAIGRKIH 60
QY 67 QNNGLDTEENNHNHNSLLAGAVAIISYLKOKLDGLKNEGLKEKIDAKKCEFTFKLEKH 126
DB 61 QNNGLDTEENNHNHNSLLAGAVAIISYLKOKLDGLKNEGLKEKIDAKKCEFTFKLEKH 120
QY 127 TDGLKEGVTDADAKAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVELTSPV 186
DB 121 TDGLKEGVTDADAKAEALITKANGTKGAEELGKLFESVEVLSKAKEMLANSVELTSPV 180
QY 187 VA 188
DB 181 VA 182

RESULT 7
ID P94234 PRELIMINARY: PRT: 193 AA.
AC P94234: 01-MAY-1997 (TREMUREL. 03, Created)
DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)
DE 01-JUN-2001 (TREMUREL. 17, Last annotation update)
OS OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSCP.
RA Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139:
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=28354;
RX MEDLINE=96296448; PubMed=8709845;
RA Lively I., Gibbs C.P., Schuster R., Dörner F.;
RT "Evidence for lateral transfer and recombination in OSCP variation in
RT Lyme disease Borrelia.";
RT Mol. Microbiol. 18:257-269(1995).
DR EMBL: I42895; AAB37003.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 20411 MW; 05B68720F061E2A0 CRC64;

Query Match 47.4%; Score 876; DB 2; Length 193;
 Best Local Similarity 100.0%; Pred. No. 3.4e-34;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 257
 DB 2 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 61
 QY 258 KIOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 317
 DB 62 KIOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 121
 QY 318 EGEHAOLGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 377
 DB 122 EGEHAOLGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 181

QY 8
 DB 122 PRELIMINARY; PRT; 194 AA.

AC 031122
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OC12;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF029871; AAB6554.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 194
 FT NON_TER 1 194
 SQ SEQUENCE 194 AA; 20640 MW; 00A5E6E2D2CE0F7F CRC64;

Query Match 47.4%; Score 876; DB 2; Length 194;
 Best Local Similarity 100.0%; Pred. No. 3.4e-34;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 257
 DB 10 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 69
 QY 258 KIOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 317
 DB 70 KIOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 129
 QY 318 EGEHAOLGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 377
 DB 130 EGEHAOLGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 189

RESULT 9
 QY 8
 DB 122 PRELIMINARY; PRT; 201 AA.

AC 09ROR8
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;

QY 198 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 257
 DB 20 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 79
 QY 258 KIOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 317
 DB 80 KIOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 139
 QY 318 EGEHAOLGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 377
 DB 140 EGEHAOLGIENVTDENAKKAILITDAKDKGAELKFKAVENLAKAKEMLANSVKEL 199

RESULT 10
 QY 8
 DB 122 PRELIMINARY; PRT; 191 AA.

AC 044726
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RA Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;
 RL "Humoral immune response to outer surface protein C of Borrelia burgdorferi in Lyme disease: role of the immunoglobulin M response in the serodiagnosis of early infection."
 DR Infect. Immun. 62:3213-3221(1994).
 DR EMBL: U08284; AAA21460.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 191
 FT NON_TER 1 191
 SQ SEQUENCE 191 AA; 20150 MW; C49A4030F0A28717 CRC64;

Query Match 46.8%; Score 863; DB 2; Length 191;
 Best Local Similarity 99.4%; Pred. No. 1.3e-33;
 Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 258
 DB 1 NSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAKEIETLLASIDELATKAIGK 60
 QY 259 IOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 318
 DB 61 IOQNGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSKKEKEIENAKKCEDFTKKL 120

QY 319 GEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAAKEMLANSVKEL 377
 DB 121 GEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAAKEMLANSVKEL 179

RESULT 11
 044999 PRELIMINARY; PRT: 177 AA.

AC 044999;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TXCH;
 MEDLINE-95286481; Pubmed-7768799;
 Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 Hansen K.;
 RT "Evolution of the Borrelia burgdorferi outer surface protein OSpC."
 RL J. Bacteriol. 177:3036-3044(1995).
 DR EMBL: X84783; CAA59254.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 177
 FT NON_TER 1 177
 SQ SEQUENCE 177 AA; 18800 MW; 156671B9164E7A2D CRC64;

Query Match Best Local Similarity 46.1%; Score 853; DB 2; Length 177;
 Matches 174; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 NTSANSADSVKGNPLTEIKKTTSDNAVLAVKEVEALLSIDETAAKIGKRIHONNG 70
 DB 1 NTSANSADSVKGNPLTEISKRTTSDNAVLAVKEVEALLSIDETAAKIGKRIHONNG 60
 QY 71 LPTENNHNSSLGAYAVISTLIKQKLDGKLNSEKLEKIDAAKCSFTTKLEKHTDGL 130
 DB 61 LPTENNHNSSLGAYAVISTLIKQKLDGKLNSEKLEKIDAAKCSFTTKLEKHTDGL 120
 QY 131 KEGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSRAAKEMLANSVKELTSPVY 187
 DB 121 KRGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSRAAKEMLANSVKELTSPVY 177

RESULT 12
 09REH7 PRELIMINARY; PRT: 178 AA.
 AC 09REH7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B31;
 MEDLINE-94075528; Pubmed-8253951;
 RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
 RT "Polymorphism in ospC gene of Borrelia burgdorferi and
 immunoreactivity of OSpC protein: implications for taxonomy and for
 use of OSpC protein as a diagnostic antigen.";
 RT J. Clin. Microbiol. 31:2570-2576(1993).

DR EMBL: X73622; CAA52001.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 178
 FT NON_TER 1 178
 SQ SEQUENCE 178 AA; 18894 MW; D619A6AC646F14EB CRC64;

Query Match Best Local Similarity 45.6%; Score 843.5; DB 2; Length 178;
 Matches 174; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 11 NTSANSADSVKGNPLTEIKKTTSDNAVLAVKEVEALLSIDETAAKIGKRIHONNG 70
 DB 1 NTSANSADSVKGNPLTEISKRTTSDNAVLAVKEVEALLSIDETAAKIGKRIHONNG 60
 QY 71 LPTENNHNSSLGAYAVISTLIKQKLDGKLNSEKLEKIDAAKCSFTTKLEKHTD-L 129
 DB 61 LPTENNHNSSLGAYAVISTLIKQKLDGKLNSEKLEKIDAAKCSFTTKLEKHTDSF 120
 QY 130 KEGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSRAAKEMLANSVKELTSPVY 187
 DB 121 KRGVTDADAKKAILKANGTKTGAEELGKLFESVEVLSRAAKEMLANSVKELTSPVY 178

RESULT 13
 031123 PRELIMINARY; PRT: 185 AA.
 AC 031123;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_Taxid=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OC13;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF029872; AAB8655.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 185
 FT NON_TER 1 185
 SQ SEQUENCE 185 AA; 19673 MW; 58D6FEE3C7769CAF CRC64;

Query Match Best Local Similarity 45.5%; Score 841; DB 2; Length 185;
 Matches 172; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 198 NNSGKDGNTSANSADSVKGNPLTEISKRTTESNAVLAVKEIETTLASIDETATKAIGK 257
 DB 10 NNSGKDGNTSANSADSVKGNPLTEISKRTTESNAVLAVKEIETTLASIDETATKAIGK 69
 QY 258 KIQOINGLAVEAGHNGTLLAGAYTISKLTQKLDGKLNSEKLEKINAKKCSDFTKKL 317
 DB 70 KIQOINGLAVEAGHNGTLLAGAYTISKLTQKLDGKLNSEKLEKINAKKCSDFTKKL 129
 QY 318 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAAKEMLANSV 373
 DB 130 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAAKEMLANSV 185

RESULT 14
 09R7B4 PRELIMINARY; PRT: 175 AA.
 AC 09R7B4;
 ID 09R7B4;
 AC 09R7B4;

DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OC Borrelia burgdorferi (Lyme disease spirochete).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;
 RX MEDLINE=97478003; PubMed=9336916;
 RA Has N.M., Postic D., Foretz M., Baranton G.;
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
 U.S.A.'";
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;
 RL Marti-Ras N., Postic D., Foretz M., Baranton G.;
 DR EMBL, U91792; AAB81889.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 1
 FT 175 175
 SQ SEQUENCE 175 AA; 18573 MW; EC059E7BD3AC3250 CRC64;

Query Match 44.8%; Score 829; DB 2; Length 175;
 Best Local Similarity 97.6%; Pred. No. 4.6e-32;
 Matches 166; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVEALLSIDEIAKA 60
 DB 6 ISCNNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVEALLSIDEIAKA 65
 QY 61 ICKKIHONNGDLTENNHNHSLAGAVTSTLIKOKLDGKNEGLKEKIDAAKCSFTFN 120
 DB 66 ICKKIHONNGDLTENNHNHSLAGAVTSTLIKOKLDGKNEGLKEKIDAAKCSFTFN 125
 QY 121 KLEKHTDGLKGEVTDADAKKAEILKANGTKTGAEELGKLFESVYLSKA 170
 DB 126 KLEKHTDGLKGEVTDADAKKAEILKANGTKTGAEELGKLFESVYLSKA 175

DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MUL;
 RX MEDLINE=95286481; PubMed=7766799;
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 Hansen K.;
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC";
 RL J. Bacteriol. 177:3036-3044(1995).
 DR EMBL, X84779; CA59250.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1 1
 FT 178 178
 SQ SEQUENCE 178 AA; 18911 MW; EE550CE48EADF1CA CRC64;

Query Match 44.6%; Score 826; DB 2; Length 178;
 Best Local Similarity 98.8%; Pred. No. 6.5e-32;
 Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 205 NTSANSADSVKGNLTETSKITESNAVLAVKEIEFTLLASIDELATKAIGKKIQONG 264
 DB 1 NTSANSADSVKGNLTETSKITESNAVLAVKEIEFTLLASIDELATKAIGKKIQONG 60
 QY 265 LAVEAGHNGTSLAGAVTTISKLITOKLDGKNEKKEIENAKKSEDFTKLGEHAOL 324
 DB 61 LAVEAGHNGTSLAGAVTTISKLITOKLDGKNEKKEIENAKKSEDFTKLGEHAOL 120
 QY 325 GIENVTDENAKKAILITDAKDKGAEELEKLFKAVENTLAKAKEMLANSVKEL 377
 DB 121 GIENVTDENAKKAILITDAKDKGAEELEKLFKAVENTLAKAKEMLANSVKEL 173

Search completed: March 18, 2002, 10:10:54
 Job time: 977 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:31 ; Search time 118.14 Seconds
(without alignments)
232.615 Million cell updates/sec

Title: US-09-596-746a-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 371

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.1101.*
1: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1804	99.7	370	22	AAB62715	Borrelia sp chimera
2	1801	99.6	394	22	AAB62730	Borrelia sp chimera
3	1559.5	86.2	369	22	AAB62716	Borrelia sp chimera
4	1556.5	86.0	393	22	AAB62731	Borrelia sp chimera
5	1533.5	84.8	378	22	AAB62712	Borrelia sp chimera
6	1529	84.5	374	22	AAB62710	Borrelia sp chimera
7	1526	84.4	398	22	AAB62728	Borrelia sp chimera
8	1524.5	84.3	401	22	AAB62738	Borrelia sp chimera
9	1489	82.3	368	22	AAB62718	Borrelia sp chimera
10	1488.5	82.3	378	22	AAB62725	Borrelia sp chimera
11	1487.5	82.2	369	22	AAB62714	Borrelia sp chimera

12	1486	82.1	392	22 AAB62734	Borrelia sp chimera
13	1480.5	81.8	393	22 AAB62732	Borrelia sp chimera
14	1480.5	81.8	401	22 AAB62733	Borrelia sp chimera
15	1289	71.3	377	22 AAB62713	Borrelia sp chimera
16	1284.5	71.0	373	22 AAB62711	Borrelia sp chimera
17	1281.5	70.8	397	22 AAB62729	Borrelia sp chimera
18	1280	70.8	400	22 AAB62739	Borrelia sp chimera
19	1244.5	68.8	367	22 AAB62719	Borrelia sp chimera
20	1241.5	68.6	391	22 AAB62735	Borrelia sp chimera
21	1225	67.7	384	22 AAB62726	Borrelia sp chimera
22	1222	67.6	408	22 AAB62737	Borrelia sp chimera
23	1199	66.3	560	22 AAB62724	Borrelia sp chimera
24	1177.5	65.1	368	22 AAB62717	Borrelia sp chimera
25	1169.5	64.6	391	22 AAB62736	Borrelia sp chimera
26	1097	60.6	410	22 AAB62740	Borrelia sp chimera
27	1096	60.6	386	22 AAB62727	Borrelia sp chimera
28	920	50.9	190	22 AAB62701	Burgdorferi ospc
29	903	49.9	193	22 AAB62704	Burgdorferi ospc
30	886	49.0	30	22 AAB62704	Burgdorferi ospc
31	885	48.9	211	22 AAB62702	Burgdorferi ospc
32	884	48.9	130	22 AAB62707	Burgdorferi ospc
33	879	48.6	192	15 AAB60888	Borrelia 257 anti
34	852	47.1	176	15 AAB60889	Borrelia 297 anti
35	837	46.3	211	20 AAB62775	Burgdorferi ospc
36	822	45.4	177	15 AAB62774	Burgdorferi ospc
37	702	38.8	466	16 AAB75739	B31 outer surface
38	701	38.8	587	16 AAB75746	B31 Osp-A/antigen
39	692	38.3	466	16 AAB75740	B31 outer surface
40	664	36.7	193	22 AAB62709	Burgdorferi ospc
41	661.5	36.6	192	22 AAB62721	Burgdorferi ospc
42	659.5	36.5	209	22 AAB62720	Burgdorferi ospc
43	658.5	36.4	192	22 AAB62703	Burgdorferi ospc
44	658.5	36.4	210	16 AAB11935	Outer surface prot
45	656.5	36.3	192	15 AAB60886	Borrelia Ip2 Ospc

ALIGNMENTS

RESULT 1
ID AAB62715 standard; Protein: 370 AA.
XX AAB62715:
AC AAB62715:
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospc protein SPQ ID NO: 34.
XX
XX Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
XX Chimeric - Borrelia sp.
XX
XX WO200078966-A1.
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000MO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UANY) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI, 2001-050113/06.
XX N-PDB; AAF29019.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -

XX Claim 43; Page 91-92; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 370 AA;

Query Match 99.7%; Score 1804; DB 22; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.6e-104;
 Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAIG 61
 1 acnmsgkdgntsansadesvkgpmlteiskrttdsnavlavkeveallsidelakaig 60
 QY 62 KIKNDGSLDNEANRNSLLAGAVTISTLTQKLSKINGSEGLKEKTAARKCSEESTK 121
 61 kkindgsldeanrnesllagaytistltqklskingseglkekaarkcseestk 120
 QY 122 LKDNHAGLIGQVTDENAKKAILKANAGKDKGYEELEKLSGLESLSKAEMLANSVK 181
 121 lkdnhagligqvtdenakkailkanaagdkgyeeleklsgleslskaemlansvk 180
 QY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEETLTLSID 241
 181 eltspvhgnnsgkdgntsansadesvkgpmlteiskrttesnavlavkeveetltsid 240
 QY 242 ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAETEKAKKC 301
 241 elakaignkikndvslndeadhngslisgaylismnlttrkisaiskdsgelekaekakkc 300
 QY 302 SEEFATLKGHEHTDLGEGVTDNNAKKAIIKTNDKTRGADLEKLPESVKNLSKAEM 361
 301 seefatlkgehtdlggyvtdnnakkailktndktrgadelkelpesvknlskaem 360
 QY 362 LTNSVKELTS 371
 361 ltnsvkeltts 370
 Db 361 ltmsvkeltts 370

RESULT 2

AK2730
 AAB62730 standard; Protein; 394 AA.

AC AAB62730;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 64.
 XX
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UYNV) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 XX

DR WPI; 2001-050113/06.
 DR N-P-SDB; AAF29034.

XX
 PT Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 to Lyme disease -
 XX

PS Claim 43; Page 128-129; 160pp; English.

XX
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX

SQ Sequence 394 AA;

Query Match 99.6%; Score 1801; DB 22; Length 394;
 Best Local Similarity 99.7%; Pred. No. 2.6e-104;
 Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAIG 61
 2 acnnskdgntsansadesvkgpmlteiskrttdsnavlavkeveallsidelakaig 61
 Db 25 scnnskdgntsansadesvkgpmlteiskrttdsnavlavkeveallsidelakaig 84
 QY 62 KIKNDGSLDNEANRNSLLAGAVTISTLTQKLSKINGSEGLKEKTAARKCSEESTK 121
 62 kkindgsldeanrnesllagaytistltqklskingseglkekaarkcseestk 121
 Db 85 kkindgsldeanrnesllagaytistltqklskingseglkekaarkcseestk 144
 QY 122 LKDNHAGLIGQVTDENAKKAILKANAGKDKGYEELEKLSGLESLSKAEMLANSVK 181
 122 lkdnhagligqvtdenakkailkanaagdkgyeeleklsgleslskaemlansvk 181
 Db 145 lkdnhagligqvtdenakkailkanaagdkgyeeleklsgleslskaemlansvk 204
 QY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEETLTLSID 241
 182 eltspvhgnnsgkdgntsansadesvkgpmlteiskrttesnavlavkeveetltsid 241
 Db 205 eltspvhgnnsgkdgntsansadesvkgpmlteiskrttesnavlavkeveetltsid 264
 QY 242 ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAETEKAKKC 301
 242 elakaignkikndvslndeadhngslisgaylismnlttrkisaiskdsgelekaekakkc 301
 Db 265 elakaignkikndvslndeadhngslisgaylismnlttrkisaiskdsgelekaekakkc 324
 QY 302 SEEFATLKGHEHTDLGEGVTDNNAKKAIIKTNDKTRGADLEKLPESVKNLSKAEM 361
 302 seefatlkgehtdlggyvtdnnakkailktndktrgadelkelpesvknlskaem 361
 Db 325 seefatlkgehtdlggyvtdnnakkailktndktrgadelkelpesvknlskaem 384
 QY 362 LTNSVKELTS 371
 362 ltnsvkeltts 371
 Db 385 ltmsvkeltts 394

RESULT 3

AKB62716
 ID AAB62716 standard; Protein; 369 AA.

AC AAB62716;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 36.
 XX
 KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX

XX (UNIV) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Datwyler RJ, Selnost G, Dyhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR MPI: 2001-050113/06.
DR N-PSDB: AAF29020.
XX
PT Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 94-95; 160pp; English.
XX
XX The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzeili or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 369 AA;
XX
Query Match 86.2%; Score 1559.5; DB 22; Length 369;
Best Local Similarity 87.5%; Fred. No. 2.2e-89;
Matches 323; Conservative 20; Mismatches 25; Indels 1; Gaps 14.
OY 2 ACNNGSGKQDNTFANGSADSVKGPNTLETISKRTTSDNAVLLAVEYALSSIDELAKAIG 61
DB 1 acnsgkgdntfanganadesvkgpnlteiskrttsdnnavllaveyeallssidelakaig 60
OY 62 KKKINDSLDNEANRNESILAGAVYITSLITQKLSKNGSEGLKRIIAAKKCSSEFSTK 121
DB 61 kkkindgsldneanrnesilagayitstlittqklaklmgsegllkheiaaakkcssefstk 120
OY 122 LKDNHAQIGCVTDEBNAKKALKLNAGKDKGVEELEKLSLSLSLAAKEMLANSVK 181
DB 121 lkdnhagigvtdenakkalklnaagkdkgvseeleklsleslskaakemlanavk 180
OY 182 ELTSPVVGNNNGKQGNFANGSADSVKGPNTLETISKRTTESNAVLLAVEYETLLTSTD 241
DB 181 eltspvvgnnsgkgnfanganadesvkgpnlteiskrttesnavllaveyettllastid 240
OY 242 ELA-FAIGKKIKNDVSLDNEADHNGSLISGAYLISNLTIKKISAIKDSGELAEIEKAKK 300
DB 241 elactelgkikqngnglaveghngtllagayltlkltdgylknscklkekienakk 300
OY 301 CSEFTAKLKGHTDLGEGVTDNNAKKAILEKTNNDKTKGADELEKLFESVKNLSKAARE 360
DB 301 csefttkrkgshagiglvtnvdenakkallitdaakdkgaalelekllfavenlakaake 360
OY 361 MLTNSVKEL 369
DB 361 mlansvkel 369
XX
XX RESULT 4
XX AAB62731
ID AAB62731 standard; Protein: 393 AA.
XX
XX AAB62731;
XX
XX 03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 66.
XX
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
XX WO200078966-A1.

```

xx 28-DEC-2000.
PD
xx 19-JUN-2000; 2000OWO-US16915.
PF
xx 18-JUN-1999; 99US-0140042.
xx
PR
xx (UNYNY ) UNIV NEW YORK STATE RES FOUND.
PA
xx (BROOK ) BROOK BIOTECHNOLOGIES INC.
PI
xx Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
xx
DR WPI: 2001-050113/06.
xx N-PSDB; AAF29035.
xx
PR Compositions of OspC polypeptides from strains of Borrelia which cause
PI Lyme disease are used to immunize animals and detect immune responses
xx to Lyme disease -
xx
PS Claim 43; Page 131; 160pp; English.
xx
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
xx
SQ Sequence 393 AA;
xx
Query Match 86.0%; Score 1556.5; DB 22; Length 393;
Best Local Similarity 87.3%; Pred. No. 3.7e-89;
Matches 322; Conservative 21; Mismatches 25; Indels 1; Gaps 1;
xx
QY 2 ACNNSGKDGONTFANSADSVKCPNLTETSKTTTDSNAVVLAVKEVEALTSIDELAKAIG 61
Db :|||||
25 scnmgskgntsaansadesvskgnltetlskttidtnavllavkeveallssidelakaiG 84
QY 62 KKIKNDGLSDNEANRNESILAGAYTITLITOKLSKLNSEGLKKERIAAKKCSSEFSTK 121
Db :|||||
85 KKikndgslsdneanrnesilagaylitstlltqklsklngseglkelaakkcssefstk 144
QY 122 LKDNHQAQIGCVTDENAKKAILKANAAGKDKGVEELEKSLGSLSSAKAKEMLANSVK 181
Db :|||||
145 lkdnhagiqgvtidenakkailkanaagkdkgveeleklsgslsslaakemlanavk 204
QY 182 ELTSTVHGNNSGKRGNTFANSADSVMCPNLTETSKTTTESNAVVLAVKEVETLLTSD 241
Db :|||||
205 eltsptvhnmgnskrgntsaansadesvsgpnltetlsktttesnavvlavkevetllastd 264
QY 242 ELA-KAIGKKIKNDVSLDNEADHNSLSGAVLIINLITRKISARVKSDELAEETIAKAK 300
Db :|||||
265 elatkaigkkikngvslaveaghnltllagaytlslklitqkldgllknekllekenak 324
QY 301 CSEEFATAIKGEHTDLGKRGVTDNNAKKAILKTTNNDKTKGADELEKLEFSVKNLSRAKE 360
Db :|||||
325 cseefatkiqgehaqgltenvctdenakkailttdaakdkgaaleleklfkavenlakaake 384
QY 361 MLTNSVKEL 369
Db :|||||
385 mlansvkel 393
xx
RESULT 5
AAB62712
ID AAB62712 standard; Protein; 378 AA.
xx
AC AAB62712;
xx
DT 03-APR-2001 (first entry)
xx
DE Borrelia sp chimeric ospC protein SEQ ID NO: 28.
xx

```

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UYNV) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI: 2001-050113/06.
 N-PSDB: AAF29016.

Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

Claim 43; Page 83-84; 160pp; English.

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 378 AA:

Query Match 84.8%; Score 1533.5; DB 22; Length 378;
 Best Local Similarity 86.3%; Pred. No. 9.5e-88;
 Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 1 MACNNGSGKGNNTSANSADSVKGNPNTETSKITDSNAVLLAVKEVALLSSIDEL-AKA 59
 DB 1 macnngskgnntsansadesvkgpnltetnktksnavllavkeveallssideiaaka 60
 QY 60 IGRKIKNDGSLDNEANRNESLAGAVTISTLTOKLSKNGSEGLKRIAAKKCSPEFS 119
 DB 61 igkkihngngldtenhngslagayastlkqkldgk-negikeidaakkcselft 119
 QY 120 TKLKNHAGLIGQVTDENAKKAILKANAAKDKGVEELKLSGLSLESISKAEMLANV 179
 DB 120 ntkkehthdlygevytdadakaalkngt-ktkgaeelgklfesevlskaakemlans 178
 QY 180 VKELTSPVHG-----NNSGKDGNTSANSADSVKGNPNTETSKITDSNAVLLAVK 231
 DB 179 vkeeltspvhaesppkpsmwnsgkgnntsansadesvkgpnltetnktksnavllavk 238
 QY 222 EVELTLLSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGEL 291
 DB 229 eveltltsidelakalgkikndvsldeadhngslisgavllisnlttkksaalkdsge 298
 QY 292 KAEIRAKKCSPEFTAKLKGHTDLGKGVTDNNAKKAITLKTNDKTKGADLEKLFESV 351
 DB 299 kaekakksceeftaklkgentdlgkgyvtdnnaakkaalktndktkgadelekifesev 358
 QY 352 KNLKRAKEMLTNSVKELTS 371
 DB 359 knlskaakemltinsvkelts 378

RESULT 6
 AAB62710
 ID AAB62710 standard; Protein; 374 AA.

AC AAB62710;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospC protein SEQ ID NO: 24.
 XX
 KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 XX
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UYNV) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI: 2001-050113/06.
 N-PSDB: AAF29014.

Compositions of OspC polypeptides from strains of Borrelia which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

Claim 43; Page 78-79; 160pp; English.

The present invention provides compositions comprising ospC proteins and chimeric ospC proteins from members of the Borrelia genus. These may be Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as vaccines against Borrelia infection, which is spread by ticks and leads to Lyme disease.

Sequence 374 AA:

Query Match 84.5%; Score 1529; DB 22; Length 374;
 Best Local Similarity 87.0%; Pred. No. 1.8e-87;
 Matches 327; Conservative 9; Mismatches 32; Indels 8; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGNPNTETSKITDSNAVLLAVKEVALLSSIDEL-AKAI 60
 DB 1 acnnskdgntsansadesvkgpnltetnktksnavllavkeveallssideiaaka 60
 QY 61 GKRKIKNDGSLDNEANRNESLAGAVTISTLTOKLSKNGSEGLKRIAAKKCSPEFT 120
 DB 61 gkkihngngldteyhnngslagayastlkqkldgk-negikeidaakkcselftn 119
 QY 121 KLNKNHAGLIGQVTDENAKKAILKANAAKDKGVEELKLSGLSLESISKAEMLANV 180
 DB 120 klnkehthdlygevytdadakaalkngt-ktkgaeelgklfesevlskaakemlansv 178
 QY 181 KELTSPVHG-----NNSGKDGNTSANSADSVKGNPNTETSKITDSNAVLLAVKEV 235
 DB 179 keltspvhaesppkpsmwnsgkgnntsansadesvkgpnltetnktksnavllavke 238
 QY 236 LKTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAVLISLITRKISAIKDSGELAEI 295
 DB 239 ltsidelakalgkikndvsldeadhngslisgavllisnlttkksaalkdsgele 298
 QY 296 EKAKKCSPEFTAKLKGHTDLGKGVTDNNAKKAITLKTNDKTKGADLEKLFESVKNLS 355
 DB 299 ekakksceeftaklkgentdlgkgyvtdnnaakkaalktndktkgadelekifesevknls 358
 QY 356 KAAKEMLTNSVKELTS 371
 DB 359 kaakemltinsvkelts 374

```

RESULT 7
AAB62728
ID AAB62728 standard; Protein: 398 AA.
XX
AC AAB62728;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 60.
XX
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PI 19-JUN-2000; 2000WO-US16915.
PR 18-JUN-1999; 99US-0140042.
PA (UANY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX MPI: 2001-050113/06.
DR N-PSDB: AAF29032.
XX
XX Compositions of ospC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 123; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 398 AA;
SQ
Query Match 84.4%; Score 1526; DB 22; Length 398;
Best Local Similarity 86.7%; Pred. No. 2.9e-87;
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;
QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTIDSNVLLAVKEVALLSIDEL-AKAI 60
DB :|||||
DB 25 scnsqkdgnstansadesvkgpnlteiskrtidsnavllavkeavallsideiaakal 84
QY 61 GKIRKNDGSLDNEANRNESILAGAVYTISTLTOKLSKLNSEGLKEKIAAACCSEFEFT 120
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 85 gkkihngngldteyhngsllagayaistlikqldgik-neglikeidaakksetftn 143
QY 121 KLRDNDHAGLTCIGVTDENAKKATILKANAAGKDCVGELEKLSGSELSAAKEMLANSV 180
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 144 klkxhtdldgkegvtdadakeaillktngt-ktkgaeeigklfsevevlskaakemlanv 202
QY 181 KELTSFVHVG-----NNSGKDGNTSANSADSVKGPNTLEISKRTITESNAVLAVERET 235
DB :|||||
DB 203 keltspvvaesppkmvmsngkdntsansadesvkgpnlteiskrtitesnavlavkevet 262
QY 236 LRTSIDELAKAIGKRTKNDVSLDNEADHNSGLTSGAVLISNLTITKRTISAIGKSGELKAET 295
DB :|||||
DB 263 lltstidelakalqkkikndvslndneadngsllsagayllsnlltltkksaikkdsgeikael 322
QY 296 EKAKKCESEFTAKLKGEBHVDLKGEGVTDNAKKAAILKTNDKTKGADLEKLEESVKNLS 355

```

```

DB 323 ekakcseeftaklkgchldlgkegvtdnakkailktndktckgadelcklfesvknls 382
QY 356 KAAKEMLTNSVKELTS 371
DB :|||||
DB 383 kaakemltnsvkelts 398
RESULT 8
AAB62738
ID AAB62738 standard; Protein: 401 AA.
XX
AC AAB62738;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 80.
XX
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PI 19-JUN-2000; 2000WO-US16915.
PR 18-JUN-1999; 99US-0140042.
PA (UANY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
PI Datwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX MPI: 2001-050113/06.
DR N-PSDB: AAF29042.
XX
XX Compositions of ospC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 150-151; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 401 AA;
SQ
Query Match 84.3%; Score 1524.5; DB 22; Length 401;
Best Local Similarity 86.0%; Pred. No. 3.7e-87;
Matches 326; Conservative 10; Mismatches 32; Indels 11; Gaps 4;
QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTIDSNVLLAVKEVALLSIDEL-AKAI 60
DB :|||||
DB 25 scnsqkdgnstansadesvkgpnlteiskrtidsnavllavkeavallsideiaakal 84
QY 61 GKIRKNDGSLDNEANRNESILAGAVYTISTLTOKLSKLNSEGLKEKIAAACCSEFEFT 120
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 85 gkkihngngldteyhngsllagayaistlikqldgik-neglikeidaakksetftn 143
QY 121 KLRDNDHAGLTCIGVTDENAKKATILKANAAGKDCVGELEKLSGSELSAAKEMLANSV 180
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 144 klkxhtdldgkegvtdadakeaillktngt-ktkgaeeigklfsevevlskaakemlanv 202
QY 181 KELTSFVHVG-----NNSGKDGNTSANSADSVKGPNTLEISKRTITESNAVLAVERE 232
DB :|||||
DB 203 keltspvvaesppkpmvmsngkdntsansadesvkgpnlteiskrtitesnavlavke 262

```

QY 233 VETLTSLIDELAKAIGKRTKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELK 292
 |||||
 Db 263 velLlSIdelAkAIGKRTKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELK 322
 QY 293 AELERAKKCSSEFTAKLGEHTDLGKEGVTDDNNAKKAIIKTNDKTGADLEKLFESVK 352
 |||||
 Db 323 aeLeKakKCSSEFTakLgEhtdlGkegVtdnakkalktndktgadelEkLfesvk 382
 QY 353 NLSKAKEMLTNSVKELT 371
 |||||
 Db 383 nlskaakemltnsvkelts 401

RESULT 9
 AAB62718
 ID AAB62718 standard; Protein: 368 AA.

AAB62718;

DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospC protein seq ID NO: 40.
 XX
 KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 XX Chimeric - Borrelia sp.
 XX
 FN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 XX
 DR WPI; 2001-050113/06.
 DR N-PSDB; AAF29022.
 XX
 PT Compositions of ospC polypeptides from strains of Borrelia which cause
 XX Lyme disease are used to immunize animals and detect immune responses
 XX to Lyme disease -
 PS Claim 43; Page 99-100; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospC proteins and
 CC chimeric ospC proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 368 AA;

QY Query Match 82.3%; Score 1489; DB 22; Length 368;
 Best Local Similarity 83.8%; Pred. No. 5.3e-85;
 Matches 311; Conservative 25; Mismatches 31; Indels 4; Gaps 4;
 QY 2 ACNNSGKDGNTSANSADSVKGPULTETSKRTITDSNVLAVKEVALSSIDELA-KAI 60
 |||||
 Db 1 acnmsgkdgntsansadesvkgpultetstskrtitesnavlavkevelssidelakai 60
 QY 61 CKKINDDSLDNEANRNSLAGAYTTSTLTQKLSKNGSEGLKERTIAAKKCSSEFT 120
 |||||
 Db 61 gkKlgnng-leangskntslisgaysldlaeklnvlknee-lkekldtakgscstefn 118
 QY 121 KIKDNHAGLGIGQVTDENAKKAIILKANAGKDKGVLELEKLSGSLSLRAAKEMLANSV 180
 |||||

Db 119 kIksehavIglndltcdnagraIlkKh-ankdkyaaeleKlIkavenIskaagdtIknav 177
 QY 181 KELTSPVHHNNSGKDGNTSANSADSVKGPULTETSKRTITDSNVLAVKEVELTSTI 240
 |||||
 Db 178 keltspIvhngmsgkdgntsansadesvkgpultetstskrtitesnavlavkeveltsi 237
 QY 241 DELAKAIGKRTKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAELEKAKK 300
 |||||
 Db 238 deLaKaIgKRTKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKaeLeKakK 297
 QY 301 CSEFTAKLGEHTDLGKEGVTDDNNAKKAIIKTNDKTGADLEKLFESVKNLKAKE 360
 |||||
 Db 298 cseftaklgehtdlgkegVtdnakkalktndktgadelEkLfesvknlskaake 357
 QY 361 MLTNSVKELT 371
 |||||
 Db 358 mltnsvkelts 368

RESULT 10
 AAB62725
 ID AAB62725 standard; Protein: 378 AA.

AAB62725;

DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospC protein seq ID NO: 54.
 XX
 KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 XX Chimeric - Borrelia sp.
 XX
 FN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 XX
 DR WPI; 2001-050113/06.
 DR N-PSDB; AAF29029.
 XX
 PT Compositions of ospC polypeptides from strains of Borrelia which cause
 XX Lyme disease are used to immunize animals and detect immune responses
 XX to Lyme disease -
 PS Claim 43; Page 114-115; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospC proteins and
 CC chimeric ospC proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 378 AA;

QY Query Match 82.3%; Score 1488.5; DB 22; Length 378;
 Best Local Similarity 83.3%; Pred. No. 5.8e-85;
 Matches 309; Conservative 31; Mismatches 28; Indels 3; Gaps 2;
 QY 1 MACNNSGKDGNTSANSADSVKGPULTETSKRTITDSNVLAVKEVALSSIDELAKAI 60
 |||||
 Db 1 macnsgkdgntsansadesvkgpultetstskrtitdsnavlavkevelssidelakai 60
 |||||

QY 61 GKKIKNDGSLDNEANRNESLAGAVTISTLTITOKLSKNGSEGIKEKIAAKKCEEST 120
 DB 61 gkikndgsldneanrnesllagavtistltitqkskngseglkekaaakcseest 120
 QY 121 KLDNHAQIGVTDENAKKAILKANAAAGDKGVEELEKLSGSELSKAEMLANSV 180
 DB 121 kldnhaqigvtdenakkaillkanaagdkgveeleklsgsleslskaekemlansv 180
 QY 181 KELTSPVHGNSGKDGNTSANSADSVKGPMLTEISKKTESNAVVLAVKEVETLTSID 240
 DB 181 keltspvhnsgnsg--gdaastnpdesakgnltvlskktltdsnafllavkevealltsid 238
 QY 241 DELAKIKGKIKNDVSLDNEADHNGSLISGAYLISNLITKKSIAIKDSGELKAEIERAKK 300
 DB 239 delakikgkikndvsltdneanrnesllagayelsklltqlslv1-nseelkkkikaekd 297
 QY 301 CSEEFATKLGEGHTDLGEGVTDNNAKAILKTNNDKTGADDELKLFESYKNTLSKAKE 360
 DB 298 csqkfttkldshaelgigsvqddnakkailkchgtkdakaleelfkleslskaaga 357
 QY 361 MLTNSVKELTS 371
 DB 358 altnsvkeltcn 368
 RESULT 11
 AAB62714
 ID AAB62714 standard; Protein: 369 AA.
 AC AAB62714;
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 32.
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 PD 28-DEC-2000.
 PF 19-JUN-2000; 2000MO-US16915.
 PR 18-JUN-1999; 99US-0140042.
 XX (UYNV) UNIV NEW YORK STATE RES FOUND.
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI: 2001-050113/06.
 DR N-PSDB: AAF29018.
 XX Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 PS Claim 43; Page 89; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzeili or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX Sequence 369 AA;
 SQ
 Query Match 82.2%; Score 1487.5; DB 22; Length 369;

Best Local Similarity 84.4%; Pred. No. 6,5e-85;
 Matches 313; Conservative 24; Mismatches 31; Indels 3; Gaps 3;
 QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKKTESNAVVLAVKEVALLSIDELAKAIG 61
 DB 1 acnnskgdntsansadsvkgnltvlskktltdsnafllavkeveallssidelaiga 60
 QY 62 KKIKNDSLDNEANRNESLAGAVTISTLTITOKLSKNGSEGIKEKIAAKKCEEST 121
 DB 61 kkinndgsldneanrnesllagavtistltitqkskngseglkekaaakcseest 120
 QY 122 LKDNHAQIGVTDENAKKAILKANAAAGDKGVEELEKLSGSELSKAEMLANSVK 181
 DB 121 lkdnhaqigvtdenakkaillkanaagdkgveeleklsgsleslskaekemlansv 180
 QY 182 ELTSPVHGNSGKDGNTSANSADSVKGPMLTEISKKTESNAVVLAVKEVETLTSID 241
 DB 181 eltspvhngnsrkdgnstnsadsvkgnltvlskktltdsnafllavkevealltsid 240
 QY 242 ELA-KAIGKIKNDVSLDNEADHNGSLISGAYLISNLITKKSIAIKDSGELKAEIERAKK 300
 DB 241 elackaigkikgn-gleangskntslsgayalsidlaeklnv1kne-elkekldtakq 298
 QY 301 CSEEFATKLGEGHTDLGEGVTDNNAKAILKTNNDKTGADDELKLFESYKNTLSKAKE 360
 DB 299 cstefcnklselhavlgldntdngaailkknankdgaalekfkavenlskaagd 358
 QY 361 MLTNSVKELTS 371
 DB 359 tlknvkelts 369

RESULT 12
 AAB62734
 ID AAB62734 standard; Protein: 392 AA.
 AC AAB62734;
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 72.
 KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 PD 28-DEC-2000.
 PF 19-JUN-2000; 2000MO-US16915.
 PR 18-JUN-1999; 99US-0140042.
 XX (UYNV) UNIV NEW YORK STATE RES FOUND.
 XX (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI: 2001-050113/06.
 DR N-PSDB: AAF29038.
 XX Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 PS Claim 43; Page 139-140; 160pp; English.
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzeili or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 392 AA;

Query Match 82.1%; Score 1486; DB 22; Length 392;
 Best Local Similarity 83.6%; Pred. No. 8,7e-85;
 Matches 310; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDELAKAI 60
 Db 25 scnsygdqnsansadesvkgpnlteiskrttesnavlavkeveallsideakai 84
 QY 61 GKTKNDGSLDNEANRNESILAGAYTITSTLTKRLSKLNGSEGLKRIIAAKKCEEFST 120
 Db 85 gkktgngn-leangskntslsagayalsdliaeklnvlknee-lkekideakgseftn 142
 QY 121 KLNKHAOLGIGVTDENAKKAILKANAGKDKGVEELEKLSGSLSKAAKEMIANSV 180
 Db 143 klkeheavlgidntdnagralikhn-andkdgaaeleklfkavenlskaadqtlknay 201
 QY 181 KELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVEETLTST 240
 Db 202 keltsplvhgnsqkdgntsansadesvkgpnlteiskrttesnavlavkeveclltsl 261
 QY 241 DELAKAIGKTKINDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELAKAIEKAK 300
 Db 262 delakaiqkktkndvsldeadngslisgaylislnttkksaisaidsgelekaieak 321
 QY 301 CSEEFYAKLKGHEHTDLGEGVTDNNAKAILKTNNDKTGADLEKLFESYKLSKAKE 360
 Db 322 cseefaklkgentdigkegvtddnakkailkcnndktgadelkelfesvknlskaake 381
 QY 361 MLTNSVKELTS 371
 Db 382 mltnsvkelts 392

RESULT 13

AAB62732 standard; Protein; 393 AA.

03-APR-2001 (first entry)

Borrelia sp chimeric ospC protein SEQ ID NO: 68.

Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PE 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29036.

Compositions of ospC polypeptides from strains of Borrelia which cause
 Lyme disease are used to immunize animals and detect immune responses
 to Lyme disease -

PS Claim 43; Page 133-134; 160pp; English.

CC The present invention provides compositions comprising ospC proteins and
 CC chimeric ospC proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

SQ Sequence 393 AA;

Query Match 82.1%; Score 1484.5; DB 22; Length 393;
 Best Local Similarity 84.1%; Pred. No. 1.1e-84;
 Matches 312; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVEALLSIDELAKAI 61
 Db 25 scnsygdqntsansadesvkgpnlteiskrttidsnavlavkeveallsideakalg 84
 QY 62 KTKNDGSLDNEANRNESILAGAYTITSTLTKRLSKLNGSEGLKRIIAAKKCEEFSTK 121
 Db 85 ktkndgsldeanrnesllagaytistltqklslkngseglkriiaakcseefstk 144
 QY 122 LKDNHAOLGIGVTDENAKKAILKANAGKDKGVEELEKLSGSLSKAAKEMIANSV 181
 Db 145 lkdnhaqlgigvtidenakkailkanaagkdkgveeleklsgslskskaakemiansv 204
 QY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAVKEVEETLTST 241
 Db 205 eltsplvhgnsrkdqntsansadesvkgpnlteiskrttesnavlavkeveclltsid 264
 QY 242 ELA-RAIKTKINDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELAKAIEKAK 300
 Db 265 elatkaigktkngn-gleangskntslsagayalsdliaeklnvlkne-elkekideakg 322
 QY 301 CSEEFYAKLKGHEHTDLGEGVTDNNAKAILKTNNDKTGADLEKLFESYKLSKAKE 360
 Db 323 cseefnkkeheavlgidntdnagralikhnandktgadelkelfesvknlskaagd 382
 QY 361 MLTNSVKELTS 371
 Db 383 tlknsvkelts 393

RESULT 14

AAB62733 standard; Protein; 401 AA.

03-APR-2001 (first entry)

Borrelia sp chimeric ospC protein SEQ ID NO: 70.

Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PE 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYNY) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29037.
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease.
XX
PS Claim 43; Page 136-137; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 401 AA:

Query Match 81.8%; Score 1480.5; DB 22; Length 401;
Best Local Similarity 83.0%; Pred. No. 26-84; Mismatches 28; Indels 3; Gaps 2;
Matches 307; Conservative 32; Mismatches 28; Indels 3; Gaps 2;

2 ACNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLAVKEVEALLSSIDELAKAIG 61
:|||||
25 scnsqkdgntsansadesvkgpnlteiskitdsnavllavkeveallssidelakaig 84
62 KRIKNGSLDNEANRNESLLAGAYTITLTQKLSKNGSEGLKRIIAAKKCEEPSTK 121
85 krlkngsl dneanr nesllagaytltltqklskngseglkriiaakkceefstk 144
122 LKDNHAGLGIQGYTDENAKKAILKANAGAKDKGVELEKLSGSLSELSKAAREMLANSVK 181
145 lkdnhaglgigvytdenakkailkanaagkdkgveeleklsgslseleskaakemlansvk 204
182 ELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLAVKEVEALLSSID 241
205 eltspvvhgnns--gdsaelnpdesakgpnlteiskitdsnavllavkeveallssid 262
242 ELAKAIKKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAKNGELKAKKCEEPSTK 301
263 elakai gkrikndvsl dneadhn gslisgayl isnltitk isakngelk akkceefstk 321
302 SEEFPAKLGEHTDLGEGVTDNNAKAILKTNDKTKGADLEKLFESYKNTLSKAEM 361
322 sefptklgehtdlgegvtdnnakailktndktkgadlek lfevykntlskaem 381
362 LTNVSKELTS 371
382 ltnsvkeltn 391

MULT 15
ID AAB62713 standard; Protein: 377 AA.
XX
AC AAB62713;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX

PA (UNYV) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
DR N-PSDB: AAF29017.
XX
CC Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease.
XX
PS Claim 43; Page 86-87; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 377 AA:

Query Match 71.3%; Score 1289; DB 22; Length 377;
Best Local Similarity 74.1%; Pred. No. 1.4e-72; Mismatches 56; Indels 12; Gaps 5;
Matches 281; Conservative 30; Mismatches 56; Indels 12; Gaps 5;

1 MACNNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLAVKEVEALLSSIDEL- 59
1 macnsgkdgntsansadesvkgpnlteinkitdsnavllavkeveallssidelatka 60
60 IGRKIRKNGSLDNEANRNESLLAGAYTITLTQKLSKNGSEGLKRIIAAKKCEEPSTK 119
61 igrkir kngsl dneanr nesllagaytltltqklskngseglkriiaakkceefstk 139
120 TKLKNHAGLGIQGYTDENAKKAILKANAGAKDKGVELEKLSGSLSELSKAAREMLANS 179
120 tklnhaglgigvytdenakkailkanaagkdkgveeleklsgslseleskaakemlans 178
180 VKELTSPVHNG-----NNSGKDGNTSANSADSVKGPNTLEISKITDSNAVLAVK 231
180 vkelts pvhng-----nns gkd gntsans adsvk gpnlteiskitdsnavlavk 238
179 vkelts pvaespkkpamvnnsgkdgntsansadesvkgpnlteiskitdsnavlavk 238
232 EVELTSLTIDELA-KAIKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAKNGE 290
232 evelts ltidela-kai gkrikndvsl dneadhn gslisgayl isnltitk isakng 296
239 etelliasidelatkaigkriqngglaveagmgtllagaytiskliltqkldglnsek 358
291 LKAETKAKKCEEPFAKLGEHTDLGEGVTDNNAKAILKTNDKTKGADLEKLFES 350
291 lkaetk akkceefptklgehtdlgegvtdnnakailktndktkgadlek lfevykntlskaem 358
299 lkekienakkcedfckkllegehaglgienytdenakkailitdaakdkgaalekifka 358
351 VKNLSKAAREMLTNSVKEL 369
351 vknlskaaremlt nsvkel 377
359 venlakaakemlansvkel 377

Search completed: March 18, 2002, 09:54:32
Job time: 330 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:33 ; Search time 55.5 Seconds
(without alignments)
150.427 Million cell updates/sec

Title: US-09-596-746A-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGTSANSADSE.....KNLSKAKEMLTNSVKELTS 371

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgcn2_6/ptodata/2/1aa/5a_COMB.pep:*
2: /cgcn2_6/ptodata/2/1aa/5a_COMB.pep:*
3: /cgcn2_6/ptodata/2/1aa/5a_COMB.pep:*
4: /cgcn2_6/ptodata/2/1aa/5a_COMB.pep:*
5: /cgcn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgcn2_6/ptodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	38.8	466	US-08-235-836C-107	Sequence 107, App
2	701	38.8	588	US-08-235-836C-122	Sequence 122, App
3	692	38.3	466	US-08-235-836C-110	Sequence 110, App
4	658.5	36.4	209	US-09-196-293-15	Sequence 15, App1
5	658.5	36.4	210	US-08-158-353-3	Sequence 3, App1
6	653.5	36.1	210	US-08-235-836C-30	Sequence 30, App1
7	652	36.0	207	US-08-235-836C-36	Sequence 36, App1
8	651.5	36.0	210	US-08-209-603E-15	Sequence 15, App1
9	597.5	33.0	212	US-08-158-353-2	Sequence 2, App1
10	588	32.5	212	US-08-031-293-2	Sequence 2, App1
11	588	32.5	212	US-07-903-580-2	Sequence 2, App1
12	556	30.7	212	US-08-158-353-4	Sequence 4, App1
13	554	30.6	209	US-08-235-836C-32	Sequence 32, App1
14	554	30.6	212	US-09-196-293-11	Sequence 11, App1
15	554	30.6	212	US-08-209-603E-11	Sequence 11, App1
16	551	30.6	212	US-08-235-836C-34	Sequence 34, App1
17	166.5	9.2	1713	US-08-600-982-24	Sequence 24, App1
18	166.5	9.2	1713	PCT-US94-10261A-24	Sequence 24, App1
19	164.5	9.1	296	US-08-320-161-11	Sequence 11, App1
20	164.5	9.1	296	US-08-137-175A-3	Sequence 3, App1
21	164.5	9.1	296	US-08-479-017-3	Sequence 3, App1
22	164.5	9.1	296	US-08-455-829-11	Sequence 11, App1
23	164.5	9.1	296	US-08-235-836C-22	Sequence 22, App1
24	163.5	9.0	2482	US-08-328-254-6	Sequence 6, App1
25	163.5	9.0	3248	US-08-353-700-1	Sequence 1, App1
26	163.5	9.0	3248	PCT-US95-16216-1	Sequence 1, App1
27	159.5	8.8	281	US-07-941-523-23	Sequence 23, App1

ALIGNMENTS

28	159.5	8.8	380	4	US-08-235-836C-118	Sequence 118, App
29	159.5	8.8	393	4	US-08-235-836C-114	Sequence 114, App
30	159.5	8.8	441	4	US-08-235-836C-120	Sequence 120, App
31	159.5	8.8	454	4	US-08-235-836C-116	Sequence 116, App
32	152.5	8.4	1786	4	US-08-973-462-8	Sequence 8, App1
33	151.5	8.4	3111	2	US-08-460-309-4	Sequence 4, App1
34	151.5	8.4	3111	2	US-08-125-077-4	Sequence 4, App1
35	143.5	7.9	688	3	US-09-141-047-8	Sequence 8, App1
36	140.5	7.8	1027	4	US-08-446-137B-2	Sequence 2, App1
37	139.5	7.7	1561	3	US-08-894-017-23	Sequence 23, App1
38	137	7.6	1147	3	US-08-470-260-5	Sequence 5, App1
39	137	7.6	1147	3	US-08-471-491-5	Sequence 5, App1
40	137	7.6	1147	4	US-08-466-662-5	Sequence 5, App1
41	137	7.6	1566	2	US-08-687-956A-23	Sequence 23, App1
42	137	7.6	3289	2	US-08-477-451-2	Sequence 2, App1
43	136	7.5	1588	5	PCT-US93-07261-11	Sequence 11, App1
44	136	7.5	1663	5	PCT-US93-07261-16	Sequence 16, App1
45	134	7.4	1196	1	US-08-144-121-4	Sequence 4, App1

RESULT 1
US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-107
Query Match 38.8%; Score 702; DB 4; Length 466;
Best Local Similarity 47.5%; Pred. No. 2.2e-43;
Matches 173; Conservative 54; Mismatches 119; Indels 18; Gaps 6;
19 ESVGPNLTETSKITDSNAVLAVKEVALLSIDELAKIKKIDSLDNEANNE 78

```

Db 100 EYFKEDGKTLVSKVTSMDKSTEEKFNEKEVESEKITTTRADGTRLEYYTGKSGSGKAK 159
QY 79 SLLAGVITSTLITOKLSK--NSEGLKERTIAAKKCEFFSTKLDKNIAQIGCVTD 136
Db 160 EYFKVLEGLTGTAKKTLTVYKEGTVLTSKNISK---SGEVSVELNDTSSAATKTTAA 215
QY 137 ENAKKAILKANAGK-----DKGYELEKLSGSEISKAKEMLANSVKELTSPVY 188
Db 216 WNDSTSTLTISADSKKTKDVLFLTDGTTIVQVYNTAGTSLGSGASET--KNLSELKNALK 273
QY 189 HGNNGKGNTSANSADSVKGPMLTEISKRTESNAVLAWEVEELLTSTDEL-AKAI 247
Db 274 MANNKGKGNTSANSADSVKGPMLTEISKRTESNAVLAWEVEELLTSTDEL-AKAI 333
QY 248 GKRIKNDVSLDNEADHNGSLISGAYLISNLITKISAIIDSGELAELEKAKKCEFTTA 307
Db 334 GKRIHONGNGLDTEYNHNGSLLAGRYAISTLIKOKLDGLKNEG-LKEKIDAAKCEFTTN 392
QY 308 KLGEHTDLKEGVTDNNAKAILKTNNDKTGADLEKLPESVKNLSKAKEMLTNSVK 367
Db 393 KLAARHTDLKEGVTDADAKAAILKTNNGTGTGAELGLTFESVVLTSKAKEMLANSVK 452
QY 368 ELTS 371
Db 453 ELTS 456

RESULT 2
US-08-235-836C-122
; Sequence 122, Application US/03235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-836C-122

```

```

Query Match 38.8%; Score 701; DB 4; Length 588;
Best Local Similarity 46.3%; Pred. No. 3,4e-43;
Matches 181; Conservative 56; Mismatches 86; Indels 68; Gaps 15;

QY 6 SGKGNSTANSADSVKGPMLTEISKRTIDSNVLAWEVEALLSIDELAKAIGKRIK 65
Db 119 SSKVTKKQGSTTEETLANKLND--SKRTFSNGTLEYEQI---TDADNATKAV-ETLK 171
QY 66 MDGSLDNEANRNESLAGAYTI-----STLIQKLSK-----LNGSEGLKERTIAAKK 113
Db 172 NSILLEG-----SLVVGTTVIRKEGIVTLKRELEKDKGVKVFINDPAGSKK----- 219
QY 114 CSEEFSTKLDKNIAQIGCVTDENAKK-----ALKANAGKDKGVEELEKL 161
Db 220 -----TGKMDSTSTLTJIA-----DSKTKDVLFLTDGTTIVQVYNTAG-----TSLEGS 265
QY 162 SGSELSLSKAKEMLANSVKELTSPVHGNNSGKGNSTANSADSVKGPMLTEISKRT 221
Db 266 ASEIKMLSE-----LKNALK--GHPM--GNNSGKGNSTANSADSVKGPMLTEISKRT 316
QY 222 ESNNAVLAWEVEETLITSIDEL-AKAIGKIKNDVSLDNEADHNGSLISGAYLISNLITK 280
Db 317 DSNNAVLAWEVEALLSIDETIAAKAIGKTIHONGNGLDTEYNHNGSLLAGRYAISTLIKQ 376
QY 281 KISAIRDSGELAELEKAKKCEFTTAKLGEHTDLKEGVTDNNAKAILKTNNDKTG 340
Db 377 KLDGLKNEG-LKEKIDAAKCEFTTNKLEKHTDLKEGVTDADAKAAILKTNNGTGTGK 435
QY 341 ADELKLPESVKNLSKAKEMLTNSVKELT 371
Db 436 ADELKLPESVKNLSKAKEMLANSVKELT 466

RESULT 3
US-08-235-836C-110
; Sequence 110, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid

```

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110

Query Match 38.3%; Score 692; DB 4; Length 466;
Best Local Similarity 48.3%; Pred. No. 1.1e-42;
Matches 185; Conservative 41; Mismatches 123; Indels 34; Gaps 11;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETISKITTDSNAVLAVKEVALLSIDEL-
AKA 59
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETISKITTDSNAVLAVKEVALLSIDEL-
AKA 76
QY 60 IGRKIKNDGSLDNEANRNESLLAGATYITSLTQKLSKLSGSEGLKRIAAKCSSEFS
119
DB 77 IGRKIHONNGLDTEYHNHNSLLAGRYAISPLIKOKLDGLK-NEGKLEKIDAAKCSSEFTF
135
QY 120 TKLKDHAOLGIGYTDENKKAAILKANAGKDKGVELEKLSGSELSKAKKEMLANS
179
DB 136 NKLLAKHTDLCKEGYTDADAKKAILKTNGT-KTKGAELKLEPESVYLSKAKKEMLANS
194
QY 180 VKELTSPVYHGNNSGKDGNTS-NSADESVKGNLTETISKITTESNA-VVLAVKEY
233
DB 195 VKELTSPVYAESPKPKQNVSVLDEKNVSVDPGEKMYLVSKKKNKDKYDIAATVDKL
254
QY 234 ETLTTS-DELAKAIGKIKNDVSLDNEADHNSLIGAVLI-SNLTTKISA
284
DB 255 ELKGTSDKNNSGVLEGVADKSKVLTJIS-DDLGOTTELVREKDEKTLVSKKVT
309
QY 285 IRDSGELKAEIEKAKKSEFTAKLGEHTDLGEGYTDONAKKAILKTNDKTKGADEL
344
DB 310 -KDKSSTEKEFNKEGEVSEKIIITRADG--TRLEYTGIKSDSGKAEVLKGYVLEGLTA
366
QY 345 EKLFESV-----NLSKAAR 359
DB 367 EKTTLVYKEGTVTLTKNISKSGE 389

RESULT 4
US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia burgdorferi
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196, 293
EARLIER FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209, 603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862, 535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 209
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match 36.4%; Score 658.5; DB 4; Length 209;

Best Local Similarity 76.7%; Pred. No. 1.1e-40;
Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETISKITTDSNAVLAVKEVALLSIDEL-
AKA 59
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETISKITTDSNAVLAVKEVALLSIDEL-
AKA 76
QY 60 IGRKIKNDGSLDNEANRNESLLAGATYITSLTQKLSKLSGSEGLKRIAAKCSSEFS
119
DB 77 IGRKIHONNGLDTEYHNHNSLLAGRYAISPLIKOKLDGLK-NEGKLEKIDAAKCSSEFTF
135
QY 120 TKLKDHAOLGIGYTDENKKAAILKANAGKDKGVELEKLSGSELSKAKKEMLANS
179
DB 136 NKLLAKHTDLCKEGYTDADAKKAILKTNGT-KTKGAELKLEPESVYLSKAKKEMLANS
194
QY 180 VKELTSPV 188
DB 195 VKELTSPV 203

RESULT 5
US-08-158-353-3
Sequence 3, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for diagnosing Early Lyme
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158, 353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-3

Query Match 36.4%; Score 658.5; DB 1; Length 210;
Best Local Similarity 76.7%; Pred. No. 1.1e-40;
Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

QY 1 MACNNSGKDGNTSANSADSVKGNLTETISKITTDSNAVLAVKEVALLSIDEL-
AKA 59
DB 17 ISCNNSGKDGNTSANSADSVKGNLTETISKITTDSNAVLAVKEVALLSIDEL-
AKA 76
QY 60 IGRKIKNDGSLDNEANRNESLLAGATYITSLTQKLSKLSGSEGLKRIAAKCSSEFS
119
DB 77 IGRKIHONNGLDTEYHNHNSLLAGRYAISPLIKOKLDGLK-NEGKLEKIDAAKCSSEFTF
135

Y 180 VKELTSPV 188

b 193 KELTNPV 200

RESULT 8
US-08-209-603E-15
; Sequence 15, Application US/08209603E
; Patent No. 6248538


```

: GENERAL INFORMATION:
: APPLICANT: FUCHS, RENATE
: APPLICANT: WILSK, BETTINA
: APPLICANT: PRAC-MURISIC, VERA
: APPLICANT: MOTZ, MANFRED
: APPLICANT: SOUTSCHECK, ERWIN
: TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
: TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
: STREET: 99 PARK AVENUE
: CITY: NEW YORK
: STATE: NY
: COUNTRY: USA
: ZIP: 10016
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" FLOPPY DISC
: COMPUTER: AT&T - IBM COMPATIBLE
: OPERATING SYSTEM: MS-DOS Version 6.2
: SOFTWARE: ASCIT
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/209,603E
: FILING DATE: 10-MAR-1994
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP90/02282
: FILING DATE: 21-DEC-1990
: APPLICATION NUMBER: US 07/862,535
: FILING DATE: 19-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: ROBINSON, WILLIAM R.
: REGISTRATION NUMBER: 27,224
: REFERENCE/DOCKET NUMBER: LKR-9217-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 697-3355
: TELEFAX: (212) 557-5635
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 210
: TYPE: AMINO ACID
: TOPOLOGY: LINEAR
: MOLECULE TYPE:
: DESCRIPTION: PROTEIN
: HYPOTHEICAL: N/A
: AMTI-SENSE: N/A
: FRAGMENT TYPE: N/A
: ORIGINAL SOURCE:
: ORGANISM: B. BURGDORFERI
: IMMEDIATE SOURCE:
: LIBRARY: DSM 5662
: POSITION IN GENOME: N/A
: FEATURE:
: IDENTIFICATION METHOD: amino acid analysis
: PUBLICATION INFORMATION: N/A
: US-08-209-603E-15

Query Match          36.0%; Score 651.5; DB 4; Length 210;
Best Local Similarity 76.2%; Pred. No. 3.4e-40;
Matches 144; Conservative 11; Mismatches 31; Indels 3; Gaps 3;

```

```

OY 180 VKELTSPV 188
DB 195 VKELTSPV 203

RESULT 9
US-08-158-353-2
: Sequence 2, Application US/08158353
: Patent No. 5620862
: GENERAL INFORMATION:
: APPLICANT: Padula, Steven J.
: TITLE OF INVENTION: Methods for diagnosing Early Lyme
: TITLE OF INVENTION: Disease
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/158,353
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Carroll, Alice O.
: REGISTRATION NUMBER: 33,542
: REFERENCE/DOCKET NUMBER: UCT93-05
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 617-861-9540
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 212 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-158-353-2

Query Match          33.0%; Score 597.5; DB 1; Length 212;
Best Local Similarity 69.1%; Pred. No. 2.8e-36;
Matches 134; Conservative 20; Mismatches 35; Indels 5; Gaps 4;

```


NUMBER OF SEQ ID NOS: 16
SOFTWARE: fastseq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match 30.6%; Score 554; DB 4; Length 212;
Best Local Similarity 67.8%; Pred. No. 3.9e-33;
Matches 124; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

QY 191 NNSGKDG-NTSANSADSVKGPNTLTKITTSNAVYLVAKVEVTLTIDELA-KAIG 248
DB 20 NNSGKVGILTSNPADESAGKPNLTETSKITDSNAFVLAVKEVELVLSIDELAKKAIG 79
QY 249 KRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAETEKAKKCESEFTAK 308
DB 80 OKIDNNNGLAALNNONGSLAGAYALSTLTETKSLKKNLEELKTEIAKAKKCESEFTAK 139
QY 309 LKGEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESYKNSKAKEMLTNSYKE 368
DB 140 LKSGHADLGKQDATDHAKAIIKTHTATDKAKKEFDLFESYEGLLAKAQAVALTNSYKE 199
QY 369 LTS 371
DB 200 LTS 202

RESULT 15

US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538

GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSTIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERMIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

LENGTH: 212
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE:
US-08-209-603E-11

Query Match 30.6%; Score 554; DB 4; Length 212;
Best Local Similarity 67.8%; Pred. No. 3.9e-33;
Matches 124; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

QY 191 NNSGKDG-NTSANSADSVKGPNTLTKITTSNAVYLVAKVEVTLTIDELA-KAIG 248
DB 20 NNSGKVGILTSNPADESAGKPNLTETSKITDSNAFVLAVKEVELVLSIDELAKKAIG 79
QY 249 KRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAETEKAKKCESEFTAK 308
DB 80 OKIDNNNGLAALNNONGSLAGAYALSTLTETKSLKKNLEELKTEIAKAKKCESEFTAK 139
QY 309 LKGEHTDLGKEGYTDNNAKKAIIKTNNDKTKGADLEKLFESYKNSKAKEMLTNSYKE 368
DB 140 LKSGHADLGKQDATDHAKAIIKTHTATDKAKKEFDLFESYEGLLAKAQAVALTNSYKE 199
QY 369 LTS 371
DB 200 LTS 202

Search completed: March 18, 2002, 09:55:34
Job time: 347 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:44 ; Search time 621.2 Seconds

(Without alignments)
165.825 Million cell updates/sec

Title: US-09-596-746A-34

Perfect score: 1809

Sequence: 1 MACNNSGKDGTSANSADES.....KNLSKAKEMLTNSVKELTS 371

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents: AA.Maln:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*

2: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*

3: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

6: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

8: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

9: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

10: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

11: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

12: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

13: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

14: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

15: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

16: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

17: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

18: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

19: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

20: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

21: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

22: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

23: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

24: /cgn2_6/ptodata/2/paa/US09_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1809	100.0	371	19	US-09-596-746A-34
2	1804	99.7	370	19	US-09-596-746-34
3	1801	99.6	394	19	US-09-596-746-64
4	1801	99.6	395	19	US-09-596-746A-64
5	1564.5	86.5	370	19	US-09-596-746A-36
6	1559.5	86.2	369	19	US-09-596-746-36
7	1556.5	86.0	393	19	US-09-596-746-66
8	1556.5	86.0	394	19	US-09-596-746A-66
9	1534	84.8	375	19	US-09-596-746A-24

10	1533.5	84.8	378	19	US-09-596-746-28	Sequence 28, Appl
11	1533.5	84.8	378	19	US-09-596-746A-28	Sequence 28, Appl
12	1529	84.5	374	19	US-09-596-746-24	Sequence 24, Appl
13	1526	84.4	398	19	US-09-596-746-60	Sequence 60, Appl
14	1524.5	84.3	399	19	US-09-596-746A-60	Sequence 60, Appl
15	1524.5	84.3	401	19	US-09-596-746-80	Sequence 80, Appl
16	1524.5	84.3	402	19	US-09-596-746A-80	Sequence 80, Appl
17	1494	82.6	369	19	US-09-596-746A-40	Sequence 40, Appl
18	1492.5	82.5	370	19	US-09-596-746A-32	Sequence 32, Appl
19	1489	82.3	368	19	US-09-596-746-40	Sequence 40, Appl
20	1488.5	82.3	378	19	US-09-596-746-54	Sequence 54, Appl
21	1488.5	82.3	378	19	US-09-596-746A-54	Sequence 54, Appl
22	1487.5	82.2	369	19	US-09-596-746-32	Sequence 32, Appl
23	1486	82.1	392	19	US-09-596-746-72	Sequence 72, Appl
24	1486	82.1	393	19	US-09-596-746A-72	Sequence 72, Appl
25	1484.5	82.1	393	19	US-09-596-746-68	Sequence 68, Appl
26	1484.5	82.1	394	19	US-09-596-746A-68	Sequence 68, Appl
27	1480.5	81.8	401	19	US-09-596-746-70	Sequence 70, Appl
28	1480.5	81.8	402	19	US-09-596-746A-70	Sequence 70, Appl
29	1289.5	71.3	374	19	US-09-596-746A-26	Sequence 26, Appl
30	1289	71.3	377	19	US-09-596-746-30	Sequence 30, Appl
31	1289	71.3	377	19	US-09-596-746A-30	Sequence 30, Appl
32	1284.5	71.0	373	19	US-09-596-746-26	Sequence 26, Appl
33	1281.5	70.8	397	19	US-09-596-746-62	Sequence 62, Appl
34	1281.5	70.8	398	19	US-09-596-746A-62	Sequence 62, Appl
35	1280	70.8	400	19	US-09-596-746-82	Sequence 82, Appl
36	1280	70.8	401	19	US-09-596-746A-82	Sequence 82, Appl
37	1249.5	69.1	368	19	US-09-596-746A-42	Sequence 42, Appl
38	1244.5	68.8	367	19	US-09-596-746-42	Sequence 42, Appl
39	1241.5	68.6	391	19	US-09-596-746-74	Sequence 74, Appl
40	1241.5	68.6	392	19	US-09-596-746A-74	Sequence 74, Appl
41	1230	68.0	385	19	US-09-596-746A-56	Sequence 56, Appl
42	1225	67.7	384	19	US-09-596-746-56	Sequence 56, Appl
43	1222	67.6	408	19	US-09-596-746A-78	Sequence 78, Appl
44	1222	67.6	409	19	US-09-596-746A-78	Sequence 78, Appl
45	1199	66.3	560	19	US-09-596-746-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1

US-09-596-746A-34

Sequence 34, Application US/09596746A

GENERAL INFORMATION:

APPLICANT: Dattwyler, Raymond J.

APPLICANT: Seino, Gerald

APPLICANT: Dykhuisen, Daniel

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001

CURRENT APPLICATION NUMBER: US/09/596,746A

CURRENT FILING DATE: 2000-06-19

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 34

LENGTH: 371

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: OSCP Chimera

US-09-596-746A-34

Query Match 100.0%; Score 1809; DB 19; Length 371;
Best local similarity 100.0%; Pred. No. 1.9e-120;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTSANSADESVKPNLTETSKRTTDSNAVLAVKVEALLSLDELAKAI 60

Db 1 MACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 61 GKIKNDGSLDNEANRNSLGAATITSTLITOKLSKNGSGKKEKIAAACEEFT 120
Db 61 GKIKNDGSLDNEANRNSLGAATITSTLITOKLSKNGSGKKEKIAAACEEFT 120
Qy 121 KIKNDHOLGIGYTDENAKKAILKANAAGKGVBELEKLSGSLESLAKAEKEMANSV 180
Db 121 KIKNDHOLGIGYTDENAKKAILKANAAGKGVBELEKLSGSLESLAKAEKEMANSV 180
Qy 181 KELTSPPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Db 181 KELTSPPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 241 DELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Db 241 DELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Qy 301 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Db 301 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Qy 361 LTNSVKEILTS 371
Db 361 LTNSVKEILTS 371

RESULT 2
US-09-596-746-34
Sequence 34 Application US/09596746
GENERAL INFORMATION:
APPLICANT: Datwyler, Raymond J.
APPLICANT: Selmost, Gerald
APPLICANT: Dykhuizen, Daniel
TITLE OF INVENTION: Groups of Borrelia burgdorferi
FILE OF INVENTION: Groups of Borrelia burgdorferi
CURRENT APPLICATION NUMBER: 2631.1002-001
PRIOR FILING DATE: 2000-06-16
SOFTWARE OF SEQ ID NOS: US 60/140,042
SEQ ID NO 34
LENGTH: 370
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE: OTHER INFORMATION: Ospe Chimera
US-09-596-746-34

Query Match
Best Local Similarity 99.7%; Score 1804; DB 19; Length 370;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2 ACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 1 ACNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Db 61 KIKNDGSLDNEANRNSLGAATITSTLITOKLSKNGSGKKEKIAAACEEFT 120
Qy 61 KIKNDGSLDNEANRNSLGAATITSTLITOKLSKNGSGKKEKIAAACEEFT 120
Db 121 KIKNDHOLGIGYTDENAKKAILKANAAGKGVBELEKLSGSLESLAKAEKEMANSV 180
Qy 121 KIKNDHOLGIGYTDENAKKAILKANAAGKGVBELEKLSGSLESLAKAEKEMANSV 180
Db 181 KELTSPPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 181 KELTSPPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Db 241 DELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Qy 241 DELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Db 301 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Qy 301 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Db 361 LTNSVKEILTS 371
Qy 361 LTNSVKEILTS 371

Page
Db 181 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 242 ELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Db 242 ELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Qy 302 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Db 302 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Qy 362 LTNSVKEILTS 371
Db 362 LTNSVKEILTS 371

RESULT 3
US-09-596-746-64
Sequence 64 Application US/09596746
GENERAL INFORMATION:
APPLICANT: Datwyler, Raymond J.
APPLICANT: Selmost, Gerald
APPLICANT: Dykhuizen, Daniel
TITLE OF INVENTION: Groups of Borrelia burgdorferi
FILE OF INVENTION: Groups of Borrelia burgdorferi
CURRENT APPLICATION NUMBER: 2631.1002-001
PRIOR FILING DATE: 2000-06-16
SOFTWARE OF SEQ ID NOS: US 60/140,042
SEQ ID NO 64
LENGTH: 394
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE: OTHER INFORMATION: Ospe Chimera
US-09-596-746-64

Query Match
Best Local Similarity 99.6%; Score 1801; DB 19; Length 394;
Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 25 SCNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 62 KIKNDGSLDNEANRNSLGAATITSTLITOKLSKNGSGKKEKIAAACEEFT 120
Db 62 KIKNDGSLDNEANRNSLGAATITSTLITOKLSKNGSGKKEKIAAACEEFT 120
Qy 122 KIKNDHOLGIGYTDENAKKAILKANAAGKGVBELEKLSGSLESLAKAEKEMANSV 180
Db 122 KIKNDHOLGIGYTDENAKKAILKANAAGKGVBELEKLSGSLESLAKAEKEMANSV 180
Qy 145 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Db 145 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 205 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Db 205 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAKEVEALLSIDELAKAI 60
Qy 242 ELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Db 242 ELAKAIGKIKNDVSLDNEADHNSLISGAYLISMLTKKISAIRKSGELAEKEMANSV 180
Qy 302 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Db 302 CSEFTAKLKEGHTDLGREGYTDNNAKAILKTNNDKTGADLEKLFESVKNLSKAEM 360
Qy 362 LTNSVKEILTS 371
Db 362 LTNSVKEILTS 371


```

RESULT 4
US-09-596-746a-64
; Sequence 64, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhutzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 395
; TYPE: PRT
; ORGANISM: OSCP Chimera
; 09-596-746a-64

Query Match          99.6%; Score 1801; DB 19; Length 395;
Best Local Similarity 99.7%; Pred. No. 7,6e-120;
Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSIDELAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSIDELAKAI 85
OY 62 KRIKNGSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAKKCSSEESTK 121
DB 86 KRIKNGSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAKKCSSEESTK 145
OY 122 LKDNHQAOLIGQYTDENAKKAILKANAAKDKVEEELKISGLESISKAEMLANSV 181
DB 146 LKDNHQAOLIGQYTDENAKKAILKANAAKDKVEEELKISGLESISKAEMLANSV 205
OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSID 241
DB 206 ELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSID 265
OY 242 ELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAKK 301
DB 266 ELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAKK 325
OY 302 SEEFPAKLKGEHTDLGEGYTDNNAKAILKTNNDKTKGADELEKLPESYKNSKAEM 361
DB 326 SEEFPAKLKGEHTDLGEGYTDNNAKAILKTNNDKTKGADELEKLPESYKNSKAEM 385
OY 362 LTNVSVELTS 371
DB 386 LTNVSVELTS 395

RESULT 5
US-09-596-746a-36
; Sequence 36, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhutzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042

```

```

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OSCP Chimera
; US-09-596-746a-36

Query Match          86.5%; Score 1564.5; DB 19; Length 370;
Best Local Similarity 87.6%; Pred. No. 3e-103;
Matches 324; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSIDELAKAI 60
DB 1 MACNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSIDELAKAI 60
OY 61 GKRKINDSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAKKCSSEEST 120
DB 61 GKRKINDSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAKKCSSEEST 120
OY 121 KLDNHAOLIGQYTDENAKKAILKANAAKDKVEEELKISGLESISKAEMLANSV 180
DB 121 KLDNHAOLIGQYTDENAKKAILKANAAKDKVEEELKISGLESISKAEMLANSV 180
OY 181 KELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSID 240
DB 181 KELTSPVHGNNSGKDGNTSANSADSVKGPNTLEISKRTSDSNVAVLLAVEVEALLSID 240
OY 241 DELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAK 299
DB 241 DELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEIERAK 300
OY 300 KCSSEFPKAKLGEHTDLGEGYTDNNAKAILKTNNDKTKGADELEKLPESYKNSKAEM 359
DB 301 KCSSEFPKAKLGEHTDLGEGYTDNNAKAILKTNNDKTKGADELEKLPESYKNSKAEM 360
OY 360 EMLTNVSVEL 369
DB 361 EMLTNVSVEL 370

RESULT 6
US-09-596-746-36
; Sequence 36, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhutzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OSCP Chimera
; US-09-596-746-36

Query Match          86.2%; Score 1559.5; DB 19; Length 369;

```

Best Local Similarity 87.5%; Pred. No. 1, 1e-102;
Matches 322; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSIDELAKAIG 61
Db 1 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSIDELAKAIG 60
OY 62 KIKINDGSLDNEANRNESLGAAYTISTLTOKLSKNGSEGLKEKTAARKCSEESTK 121
Db 61 KIKINDGSLDNEANRNESLGAAYTISTLTOKLSKNGSEGLKEKTAARKCSEESTK 120
OY 122 LKDNHAOLGIGVTDENAKKAILKANAAGDKVGELEKLSGLESLSKAKKEMLANSVK 181
Db 121 LKDNHAOLGIGVTDENAKKAILKANAAGDKVGELEKLSGLESLSKAKKEMLANSVK 180
OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSID 241
Db 181 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSID 240
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISMLITFKISAIRKDSGELKAIEKAK 300
Db 241 ELATKAIGKKIQONGGLAVAGHNGTLLAGAYTISTLTOKLDGLKNSKLEKIEKAK 300
OY 301 CSEEFPAKLAGEHTDLGKEGYTDNNAKAILKTNDKTGADELKLFESVKNLSKAKE 360
Db 301 CSEDFTKKLEGEHAOLGIEVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 360
OY 361 MLTNSVKEL 369
Db 361 MLANSVKEL 369
```

RESULT 7

```
US-09-596-746-66
; Sequence 66, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 393
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-66
```

Query Match 86.0%; Score 1556.5; DB 19; Length 393;
Best Local Similarity 87.3%; Pred. No. 2e-102;
Matches 322; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSIDELAKAIG 61
Db 25 SCNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSIDELAKAIG 84
OY 62 KIKINDGSLDNEANRNESLGAAYTISTLTOKLSKNGSEGLKEKTAARKCSEESTK 121
Db 85 KIKINDGSLDNEANRNESLGAAYTISTLTOKLSKNGSEGLKEKTAARKCSEESTK 144
OY 122 LKDNHAOLGIGVTDENAKKAILKANAAGDKVGELEKLSGLESLSKAKKEMLANSVK 181
Db 145 LKDNHAOLGIGVTDENAKKAILKANAAGDKVGELEKLSGLESLSKAKKEMLANSVK 204
```

```
OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSID 241
Db 205 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSID 264
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISMLITFKISAIRKDSGELKAIEKAK 300
Db 265 ELATKAIGKKIQONGGLAVAGHNGTLLAGAYTISTLTOKLDGLKNSKLEKIEKAK 324
OY 301 CSEEFPAKLAGEHTDLGKEGYTDNNAKAILKTNDKTGADELKLFESVKNLSKAKE 360
Db 325 CSEDFTKKLEGEHAOLGIEVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 384
OY 361 MLTNSVKEL 369
Db 385 MLANSVKEL 393
```

RESULT 8

```
US-09-596-746A-66
; Sequence 66, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 394
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-66
```

Query Match 86.0%; Score 1556.5; DB 19; Length 394;
Best Local Similarity 87.3%; Pred. No. 2e-102;
Matches 322; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSIDELAKAIG 61
Db 26 SCNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSIDELAKAIG 85
OY 62 KIKINDGSLDNEANRNESLGAAYTISTLTOKLSKNGSEGLKEKTAARKCSEESTK 121
Db 86 KIKINDGSLDNEANRNESLGAAYTISTLTOKLSKNGSEGLKEKTAARKCSEESTK 145
OY 122 LKDNHAOLGIGVTDENAKKAILKANAAGDKVGELEKLSGLESLSKAKKEMLANSVK 181
Db 146 LKDNHAOLGIGVTDENAKKAILKANAAGDKVGELEKLSGLESLSKAKKEMLANSVK 205
OY 182 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSID 241
Db 206 ELTSPVHGNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLLAVKEVEALLSID 265
OY 242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISMLITFKISAIRKDSGELKAIEKAK 300
Db 266 ELATKAIGKKIQONGGLAVAGHNGTLLAGAYTISTLTOKLDGLKNSKLEKIEKAK 325
OY 301 CSEEFPAKLAGEHTDLGKEGYTDNNAKAILKTNDKTGADELKLFESVKNLSKAKE 360
Db 326 CSEDFTKKLEGEHAOLGIEVTDENAKKAILITDAKDKGAELKLFKAVENLAKAKE 385
OY 361 MLTNSVKEL 369
Db 386 MLANSVKEL 394
```

RESULT 9
US-09-596-746a-24
; Sequence 24, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-24

Query Match 84.8%; Score 1534; DB 19; Length 375;
Best Local Similarity 87.0%; Pred. No. 7.5e-101;
Matches 328; Conservative 9; Mismatches 32; Indels 8; Gaps 4;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTDSNAVLAVKEVEALLSIDEL-
DB 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTDSNAVLAVKEVEALLSIDELIAKA 60
QY 60 ICKRIKNDGSLDNEANRNSLLAGAYTISTLTOKLSKNGSEGLKRIAAKCSFEFS 119
DB 61 ICKRIHONGNLDTEYNNHNSLLAGAYISTLIKOKLDGLK-NEGLEKIDAAKCSFEFT 119
QY 120 TKLKNHQAOLIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLSKAKKEMLANS 179
DB 120 NKLKEKHTDLGKEGYTDADAKKAILKTNGT-KTKGAELGLKLFESYEVLSKAKKEMLANS 178
QY 180 VKELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVE 234
DB 179 VKELTSPVVAESPAMVNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVE 238
QY 235 TLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLTTKISAIKDSGELKAE 294
DB 239 TLTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLTTKISAIKDSGELKAE 298
QY 295 IEKAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDTKGADELEKLFESYKNL 354
DB 299 IEKAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDTKGADELEKLFESYKNL 358
QY 355 SKAKKEMLTNSVKELTNS 371
DB 359 SKAKKEMLTNSVKELTNS 375

RESULT 10
US-09-596-746-28
; Sequence 28, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-28

Query Match 84.8%; Score 1533.5; DB 19; Length 378;
Best Local Similarity 86.3%; Pred. No. 8.3e-101;
Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTDSNAVLAVKEVEALLSIDEL-
DB 1 MACNSGKDGNTSANSADSVKGPMLTEISKRTDSNAVLAVKEVEALLSIDELIAKA 60
QY 60 ICKRIKNDGSLDNEANRNSLLAGAYTISTLTOKLSKNGSEGLKRIAAKCSFEFS 119
DB 61 ICKRIHONGNLDTEYNNHNSLLAGAYISTLIKOKLDGLK-NEGLEKIDAAKCSFEFT 119
QY 120 TKLKNHQAOLIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLSKAKKEMLANS 179
DB 120 NKLKEKHTDLGKEGYTDADAKKAILKTNGT-KTKGAELGLKLFESYEVLSKAKKEMLANS 178
QY 180 VKELTSPVYHG-----NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVK 231
DB 179 VKELTSPVVAESPAPKPMVNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVK 238
QY 232 EVELTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLTTKISAIKDSGEL 291
DB 239 EVELTSTIDELAKAIGKRIKNDVSLDNEADHNSLISGAYLISNLTTKISAIKDSGEL 298
QY 292 KAEIERAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDTKGADELEKLFESY 351
DB 299 KAEIERAKKCSFEFTAKLKEGVTDDNNAKKAITKTNNDTKGADELEKLFESY 358
QY 352 KNLKAKKEMLTNSVKELTNS 371
DB 359 KNLKAKKEMLTNSVKELTNS 378

RESULT 11
US-09-596-746a-28
; Sequence 28, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selinost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera

US-09-596-746a-28

Query Match 84.8%; Score 1533.5; DB 19; Length 378;
Best Local Similarity 86.3%; Pred. No. 8.3e-101;
Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

QY 1 MACNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEL-AKA 59
DB 1 MACNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEL-AKA 60
QY 60 IGKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 119
DB 61 IGKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 119
QY 120 TLKDNHAGLGIOGYTDENAKKAILKANAGKDKGVEELEKLSGSLSKAKKEMLANS 179
DB 120 NKLKKEHTDLGEGYTDADAKKAILKANGT-KTKGAEELGKLFESVEVLSKAKKEMLANS 178
QY 180 VELTSPVYHG-----NNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVK 231
DB 179 VELTSPVYAESPKRPSMVNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVK 238
QY 232 EVELTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGEL 291
DB 239 EVELTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGEL 298
QY 292 KAEIEKAKKCESEFTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADLEKLFESV 351
DB 299 KAEIEKAKKCESEFTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADLEKLFESV 358
QY 352 KNLKRAKEMLTNSVELTS 371
DB 359 KNLKRAKEMLTNSVELTS 373

RESULT 12
US-09-596-746-24
Sequence 24, Application US/09596746

GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinoost, Gerald
APPLICANT: Dykhulzen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia atzei That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 374
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OSpC Chimera
US-09-596-746-24

Query Match 84.5%; Score 1529; DB 19; Length 374;
Best Local Similarity 87.0%; Pred. No. 1.7e-100;
Matches 327; Conservative 9; Mismatches 32; Indels 8; Gaps 4;

QY 2 ACNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEL-AKAI 60
DB 1 ACNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEL-AKAI 60
QY 61 GKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 120
DB 61 GKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 120

DB 61 GKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 119
QY 121 KLKDNHAGLGIOGYTDENAKKAILKANAGKDKGVEELEKLSGSLSKAKKEMLANSV 180
DB 120 KLKKEHTDLGEGYTDADAKKAILKTNGT-KTKGAEELGKLFESVEVLSKAKKEMLANSV 178
QY 181 KELTSPVYHG-----NNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 235
DB 179 KELTSPVYAESPKRPSMVNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 238
QY 236 LTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAEI 295
DB 239 LTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAEI 298
QY 296 EKAKKCESEFTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADLEKLFESVKNLS 355
DB 299 EKAKKCESEFTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADLEKLFESVKNLS 358
QY 356 KAKKEMLTNSVELTS 371
DB 359 KAKKEMLTNSVELTS 374

RESULT 13
US-09-596-746-60
Sequence 60, Application US/09596746

GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinoost, Gerald
APPLICANT: Dykhulzen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia atzei That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 398
TYPE: PRT
ORGANISM: OSpC Chimera
US-09-596-746-60

Query Match 84.4%; Score 1526; DB 19; Length 398;
Best Local Similarity 86.7%; Pred. No. 3.1e-100;
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;

QY 2 ACNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEL-AKAI 60
DB 25 SCNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEL-AKAI 84
QY 61 GKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 120
DB 85 GKRIKNGSLDNEANRNESLAGAYTISTLTOKLSKNGSEGLKESKIAAKKCESEFT 143
QY 121 KLKDNHAGLGIOGYTDENAKKAILKANAGKDKGVEELEKLSGSLSKAKKEMLANSV 180
DB 144 KLKKEHTDLGEGYTDADAKKAILKTNGT-KTKGAEELGKLFESVEVLSKAKKEMLANSV 202
QY 181 KELTSPVYHG-----NNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 235
DB 203 KELTSPVYAESPKRPSMVNNSGDKGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVE 262
QY 236 LTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAEI 295
DB 263 LTSTIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTITKISAIKDSGELKAEI 322
QY 296 EKAKKCESEFTAKLKGHTDLGEGYTDNNAKAILKTNNDKTGADLEKLFESVKNLS 355

```

Db 323 EKAKKCEFTAKLGEHTDLCGEVTDNNAKAILKTNNDKTKGADELKLFESVKNLS 382
QY 356 KAKKEMLTNSVKELTS 371
Db 383 KAKKEMLTNSVKELTS 398

```

RESULT 14

```

US-09-596-746a-60
; Sequence 60, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Dantel
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 399
; TYPE: PRT
; ORGANISM: ospC Chlmera
US-09-596-746a-60

```

Query Match 84.4%; Score 1526; DB 19; Length 399;
 Best Local Similarity 86.7%; Pred. No. 3.1e-100;
 Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;

```

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTDTSNNAVLAVKEVEALLSSIDEL-AKAI 60
Db 26 SCNNSGKDGNTSANSADSVKGPNTLEISKRTDTSNNAVLAVKEVEALLSSIDELAKAI 85
QY 61 GKRIKIDGSLDNEANRNESLAGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPST 120
Db 86 GKRIHONNGLDTEYNHNGSLAGAVTISTLTQKLOGLK-NEGKLEKIDAKKCEFTFN 144
QY 121 KLDNHAOLGIGVTDENAKKAILKANAAGKDVGELEKLSGLSLSKRAKEMLANSV 180
Db 145 KLEKHTDLCGEVTDADAKKAILKTNGT-KTKGAELGLKLFESVEVLSKRAKEMLANSV 203
QY 181 KELTSPPVHG-----NNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAKEVEET 235
Db 204 KELTSPPVHESPAVNNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAKEVEET 263
QY 236 LKTSIDELAKAIKRTKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEI 295
Db 264 LKTSIDELAKAIKRTKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELKAEI 323
QY 296 EKAKKCEFTAKLGEHTDLCGEVTDNNAKAILKTNNDKTKGADELKLFESVKNLS 355
Db 324 EKAKKCEFTAKLGEHTDLCGEVTDNNAKAILKTNNDKTKGADELKLFESVKNLS 383
QY 356 KAKKEMLTNSVKELTS 371
Db 384 KAKKEMLTNSVKELTS 399

```

RESULT 15
 US-09-596-746-80
 ; Sequence 80, Application US/09596746
 ; GENERAL INFORMATION:
 ; APPLICANT: Dattwyler, Raymond J.
 ; APPLICANT: Selmost, Gerald
 ; APPLICANT: Dykhuitzen, Dantel

```

; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRT
; ORGANISM: ospC Chlmera
US-09-596-746-80

```

Query Match 84.3%; Score 1524.5; DB 19; Length 401;
 Best Local Similarity 86.0%; Pred. No. 4e-100;
 Matches 326; Conservative 10; Mismatches 32; Indels 11; Gaps 4;

```

QY 2 ACNNSGKDGNTSANSADSVKGPNTLEISKRTDTSNNAVLAVKEVEALLSSIDEL-AKAI 60
Db 25 SCNNSGKDGNTSANSADSVKGPNTLEISKRTDTSNNAVLAVKEVEALLSSIDELAKAI 84
QY 61 GKRIKIDGSLDNEANRNESLAGAVTISTLTQKLSKLNSEGLKEKIAAKKCEEPST 120
Db 85 GKRIHONNGLDTEYNHNGSLAGAVTISTLTQKLOGLK-NEGKLEKIDAKKCEFTFN 143
QY 121 KLDNHAOLGIGVTDENAKKAILKANAAGKDVGELEKLSGLSLSKRAKEMLANSV 180
Db 144 KLEKHTDLCGEVTDADAKKAILKTNGT-KTKGAELGLKLFESVEVLSKRAKEMLANSV 202
QY 181 KELTSPPVHG-----NNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAKE 232
Db 203 KELTSPPVHESPAVNNNSGKDGNTSANSADSVKGPNTLEISKRTTESNAVLAKE 262
QY 233 VETLLTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELK 292
Db 263 VETLLTSIDELAKAIGKRIKNDVSLDNEADHNGSLISGAYLISNLTTRKISAIKDSGELK 322
QY 293 AETEKAKKCEFTAKLGEHTDLCGEVTDNNAKAILKTNNDKTKGADELKLFESV 352
Db 323 AETEKAKKCEFTAKLGEHTDLCGEVTDNNAKAILKTNNDKTKGADELKLFESV 382
QY 353 NLSKAKEMLTNSVKELTS 371
Db 383 NLSKAKEMLTNSVKELTS 401

```

Search completed: March 18, 2002, 10:08:45
 Job time: 973 sec


```
Sequence 3, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSpC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09/974,992
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 207
TYPE: PRT
ORGANISM: Borrelia garinii
US-09-974-992-3
```

```
Query Match
Best Local Similarity 75.0%; Score 656; DB 6; Length 207;
Matches 141; Conservative 20; Mismatches 23; Indels 4; Gaps 3;

QY 1 MACNSKCGDNTSANSADSVKGNLTKITSDSNAVLAKEVEALISDELAKAI 60
Db 17 ISCNSSG--DSDASTNDESAKGNLTVISKITDSNAFLAKVEALLSIDELSKAI 74
QY 61 GKIKINGSLDNEANRNSLHAGVITSLITOKLSKNGSEGLKRIAAKCCSEFST 120
Db 75 GKIKINGSLDNEANRNSLHAGVITSLITOKLSKNGSEGLKRIAAKCCSEFST 133
QY 121 KLNDNHAQLGIVTDENAKKAILKANAAKDKGVELEKLSGLESLSKAAKEMLV 180
Db 134 KLDSHAELGIVODNNAKAILKTKHGT-KDKGAKELERKLSLESLSKAAQALNTSV 192
QY 181 KELTSPIV 188
Db 193 KELTSPIV 200
```

```
RESULT 3
US-09-974-992-7
Sequence 7, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSpC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09/974,992
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia afzelii
US-09-974-992-7
```

Query Match 30.7%; Score 556; DB 6; Length 212;
Best Local Similarity 67.2%; Pred. No. 2,9e-31;
Matches 133; Conservative 19; Mismatches 39; Indels 2; Gaps 2;

```
QY 191 NSKGDGNT--SANSADSVKGNLTKITSDSNAVLAKEVEALISDELAKAI 248
Db 20 NSKGDGDSASTNPADSDAGKPNLTKITSDSNAVLAKEVEALISDELAKAI 79
QY 249 KIKINDVSLDNEADHNSLISGAYLISNLTTRKISAIKNGSELKAIKAKCCSEFTAK 308
Db 80 QKIDNNNGLAALNNQNSLHAGVITSLITOKLSKNGSEGLKRIAAKCCSEFTAK 139
QY 309 LKGEHNDIGESTVDNNAKAILKTNNDTKGADLEKLFESYKLSKAEMLTNSYKE 368
Db 140 LKSGHADLGKQDTPDHANAAILKTHATDKGAKERKDFESYEGILKAQYALNTSYKE 199
QY 369 LTS 371
Db 200 LTS 202
```

```
RESULT 4
US-09-708-427-15045
Sequence 15045, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. Alexandrov et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15045
LENGTH: 1304
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..1304
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc-feature
LOCATION: 1..1304
OTHER INFORMATION: Ceres Seq. ID 1828628
US-09-708-427-15045
```

Query Match 10.4%; Score 187.5; DB 6; Length 1304;
Best Local Similarity 23.4%; Pred. No. 7.3e-05;
Matches 102; Conservative 75; Mismatches 149; Indels 109; Gaps 17;

```
QY 4 NSKGDGNTSANSADSVKGNLTKITSDSNAVLAKEVEALISDELAKAI 48
Db 85 NEKIKELAAOKRAESESFEYERFRAVELQAGLEAVQKKDVTSKNELSIRQHALDLSA 144
QY 49 LLSIDELAKAIKIKINGSLDNEA--NNESLHAGVITSLITOK-----LSKNG 100
Db 145 LLSTTEEL-----QKVKHLSMTADAKKNAKLSHAEPATKIAIIEHKEIILASELGRKA 199
QY 101 SEGILKRIAAKCCSEFTSTKLNDNHAQLGIVTDENAKKAILKANAAKDKGVEELE- 159
Db 200 ILGSKREKRAIE--GNEIYSKLKSE-----ILLNGELKYSILSSLSKDEGLVEQLKV 252
QY 160 --KLGSLSLSKAAKEMLVANSVKELTSPVYHNGNSKGDGNTSANSADSVKGNLTKIT 217
Db 253 DLEAAKMAESCTNSVVEEMKKNVHLEKEV-----RESNRSKSSASBSME-----SYM 300
QY 218 KRITESNAV-----LAKVEETLITSLD-----LAKAIGKIKNDV- 255
Db 301 KQLAELNHLVHETKSDNAQKEKIELEKTEIAORTDLEEGROYCIAKESASLENLVE 360
QY 256 -----SLDNEADHNSLISGAYLISNLTTRKISAIKNGSELKAIKAKCCSEFTAK 300
Db 361 STKSELISOEKTRALDNEKAATSN-----IONLLDQRTSLSELECKYEEERSKK 413
QY 301 CSEFTAKLGEHNDIGESTVDNNAKAILKTNNDTKGADLEKLFESYKLSKAEMLTNSYKE 360
```


Db 414 DMSLTAL-----QEASTESSEAKATLVQOEELKNC---ESQVDSLKLAKSETNE 462
 QY 361 ----MTNSVKELTS 371
 Db 463 KYEKMLDARNEIDS 477

RESULT 5
 US-09-708-427-15044
 ; Sequence 15044, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09
 ; NUMBER OF SEQ ID NOS: 85364
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15044
 ; LENGTH: 1313
 ; TYPE: PR
 ; ORGANISM: Arabidopsis thaliana

FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: 1..1313
 ; OTHER INFORMATION: Xaa is any amino acid
 ; NAME/KEY: misc.feature
 ; LOCATION: 1..1313
 ; OTHER INFORMATION: Ceres Seq. ID 1828627
 US-09-708-427-15044

Query Match 10.4%; Score 187.5; DB 6; Length 1313;
 Best Local Similarity 23.4%; Pred. No. 7.3e-05;
 Matches 102; Conservative 75; Mismatches 149; Indels 109; Gaps 17;

QY 4 NNSGKGNSTANSANADES-----VKPNLTETSKRTTDSNAVLAVK-----EVEA 48
 Db 94 NEKLALAAQKRAESESFEYERFAVELEQAGLEAVOKDVTSKNELESIRSHALDISA 153
 QY 49 LSSSIDELAKAIKKIKNDGSLDNEA-NRNESSLAGAVTISTLITOK-----LSKLNG 100
 Db 154 LSTTEL-----QRYKHELSMTADAKNKALSHAEETKTAETIAEAEAILASLGRLKA 208
 QY 101 SEGKERRIAAKKCSSEFTSKLDNHAQIGVTDENAKKAILKANAAKDKGVEELE- 159
 Db 209 LLSGKEEKEAIE--GNEIYVSKLSE-----IELRLGELEKVSILSESLKQEGLEVQLKV 261
 QY 160 --KLSSLESLSRAAKEMLANSVKELTSPYVHGNNSGKDGNTSANSADSEYKGNLTETIS 217
 Db 262 DLEAAKRAESCTNSVSEWKNKVHELEKEV-----EESNRKSSKSESE-----SYM 309
 QY 218 KRITESNAVY-----LAVKEVELTLTIDE-----LAKAIGKKIRNDV- 255
 Db 310 KQALAEINHVIHERKSDNAQKEKIELLEKTIIEAQRDLEEGYQVCLAKEEASLLENLVE 369
 QY 256 -----SLDNEADHNGSLISGAYLISNLITKKAISAKDSEGLAEIERAKK 300
 Db 370 SIKSELEISOEERTRALDNEKATSN-----IQNLDRTELISLEKCKVEEESKK 422
 QY 301 CSSEFTAKLGEHTDLGKEVTDNNAKAILKTNNOKTKADELEKLFESYKNSKAKE 360
 Db 423 DMSLTAL-----QEASTESSEAKATLVQOEELKNC---ESQVDSLKLAKSETNE 471
 QY 361 ----MTNSVKELTS 371
 Db 472 KYEKMLDARNEIDS 486

RESULT 6
 US-09-815-242-5883
 ; Sequence 5883, Application US/09815242

GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlssen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5883
 ; LENGTH: 837
 ; TYPE: PR
 ; ORGANISM: Staphylococcus aureus

Query Match 10.38%; Score 187; DB 6; Length 837;
 Best Local Similarity 22.28%; Pred. No. 4.4e-05;
 Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

QY 11 NTSANSADSEYKGNLTETSKRTDS-----NAVLLAVEEVALSSIDELAKAIKKIK 65
 Db 278 NTAFTNLKNGIQDONTIKQVNFDTADAKKRNATYNTQAEQIIN-----KAGPPTS 331
 QY 66 NDG-----SLDNEARNRESLAGAVTISTLITOKLSKNGS--EGLKEKJAAARKC 114
 Db 332 KDVETALLENVQRAKKNELNGN--VANAQKTAKNALNLTGINNAQKREALKSQIEGATTV 390
 QY 115 S--EESFTSKLD--NHAQIGI--GYTDENAKKAILKANAAKDKGVEELEKIS-----G 163
 Db 391 AGVAVQVSTJASLNTAMSNLQNGINDEAATRAAL-----NGTONLEKAKOHANTAI 442
 QY 164 SLESLSRAAKEMLANSVKELTSPYVHGNNSGKDGNTSA-----NSADESY--KGNLT 214
 Db 443 GLSHLITNAQKALQOLVQOSTTVAEAGQNBQKANNVDAAMKLSQIDANNTQONQYTT 502
 QY 215 EIS--KRITESNAVY-----LAVKEVELTLTIDE-----LAKAIGKKIRNDV- 255
 Db 503 DASONKKDAYNNAVTAAQIITDPTLPTVIINQAGOVSTTKNALNGENLEAAKQ 562
 QY 233 VETLITSIDELAKAIKKIKNDVSLDNEADHNGSLISGAYLISNLITKKAISAKDSEGL 292
 Db 563 ASQSLGSLDNNNAQKQVTVQINGAHVTDANDQKQANQNLNTAMGLKQALADKATK 622
 QY 293 AEI-----EKAKK-----CSSEFTAKLGEHTDLG--KEGYTDNNAKAILKTNNDK 337
 Db 623 ATVNFDTADAKKQAYNTAYNTAENAIITISKANGAQTQAEVQDAIKOVAAQALNGNANY 682
 QY 338 TKGADELEKLFESYKNSKAKEMLTNSVKELTS 371
 Db 683 QHARDEATALINSSNDLNOAQKDALQOVQNAVTT 716

RESULT 7

US-10-072-851-5883

; Sequence 5883, Application US/10072851

; GENERAL INFORMATION:

; APPLICANT: Carr, Grant J.

; APPLICANT: Xu, H. Howard

; APPLICANT: Foulkes, J. Gordon

; APPLICANT: Zamudio, Carlos

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Roemer, Terry

; APPLICANT: Jiang, Bo

; APPLICANT: Boone, Charles

; APPLICANT: Bussey, Howard

; APPLICANT: Bussey, Howard

; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits

; TITLE OF INVENTION: Proliferation

; FILE REFERENCE: ELITRA.028A

; CURRENT APPLICATION NUMBER: US/10/072,851

; CURRENT FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; NUMBER OF SEQ ID NOS: 15811

; SOFTWARE: FastSeq for Windows version 4.0

; SEQ ID NO 5883

; LENGTH: 837

; TYPE: PRN

; ORGANISM: Staphylococcus aureus

US-10-072-851-5883

Query Match 10.3%; Score 187; DB 7; Length 837;

Best Local Similarity 22.2%; Pred. No. 4.4e-05;

Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

```
QY 11 NTSANSADSVKGNPTEISKTTDS-----NAVLLAVEVEALLSIDELAKAICKK 65
DB 278 NTAAMTLKNGIDQNTIKGVNFTDADAKRNAYTNVAVQAEOILN-----KAGPNTS 331
QY 66 NCG-----SLDNEARNESILAGAYTISLITOKSLKNGS--EGIKETIAAKKC 114
DB 332 KCGVETALENVRKAKNELNGON--VANAKTTAKNALNLTSINNAOKKALQSIEGATTV 390
QY 115 S--EEFSTKLKD--NHAOLGIG--GVTDENAKKALKANAAKDKGVELEKLS-----G 163
DB 391 AGVNOVSTASELNTAMSLNONGINDEAATKAL-----NGTONLEKAKOHANTAI 442
QY 164 SLESLSKAKKEMLANVKELTSPVHGNSGKGNNTSA-----NSADESV--KGPNTL 214
DB 443 GSHLTNNAOKKALQVOOSTVVAEAGNEOKANNVDAAMDKLROSIADNATTKQONNT 502
QY 215 EIS--KITESNAVV-----LAVKE 232
DB 503 DASQNKDAVNNVNTAAGTIIDQTSPTLDPYINQAAGOVSTTKNALNGENMLEAKAQ 562
QY 233 VETLTSIDELAKAICKKIKNDVSLDNEADHNGSLISGAYLLSNLTKKISAIKDSGEIK 292
DB 563 ASQSLGSLDNLNNAOKQOYTDQINGAHYVDEANQIKONQNLMTAGNKLQAATADKATK 622
QY 293 AEI-----EKAK-----CSEFTAKIKGHTDLG--KEGVTDNNAKALKITNNDK 337
DB 623 ATVNETDADQAKOAVNTAVTNAENIISKANGSNATQAEVEQAIKOYNNAKQALNGNANV 662
QY 338 TKGADELEKLFESVKNLSKAKKEMLTNSVYELTS 371
DB 683 QHAKDEATLINSNDLNOAQKALNQOYQVQNTT 716
```

RESULT 8

US-09-815-242-13080

; Sequence 13080, Application US/09815242

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes In

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows version 4.0

; SEQ ID NO 13080

; LENGTH: 875

; TYPE: PRN

; ORGANISM: Staphylococcus aureus

US-09-815-242-13080

Query Match 10.3%; Score 187; DB 6; Length 875;

Best Local Similarity 22.2%; Pred. No. 4.6e-05;

Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

```
QY 11 NTSANSADSVKGNPTEISKTTDS-----NAVLLAVEVEALLSIDELAKAICKK 65
DB 315 NTAAMTLKNGIDQNTIKGVNFTDADAKRNAYTNVAVQAEOILN-----KAGPNTS 368
QY 66 NCG-----SLDNEARNESILAGAYTISLITOKSLKNGS--EGIKETIAAKKC 114
DB 369 KCGVETALENVRKAKNELNGON--VANAKTTAKNALNLTSINNAOKKALQSIEGATTV 427
QY 115 S--EEFSTKLKD--NHAOLGIG--GVTDENAKKALKANAAKDKGVELEKLS-----G 163
DB 428 AGVNOVSTASELNTAMSLNONGINDEAATKAL-----NGTONLEKAKOHANTAI 479
QY 164 SLESLSKAKKEMLANVKELTSPVHGNSGKGNNTSA-----NSADESV--KGPNTL 214
DB 480 GSHLTNNAOKKALQVOOSTVVAEAGNEOKANNVDAAMDKLROSIADNATTKQONNT 539
QY 215 EIS--KITESNAVV-----LAVKE 232
DB 540 DASQNKDAVNNVNTAAGTIIDQTSPTLDPYINQAAGOVSTTKNALNGENMLEAKAQ 599
QY 233 VETLTSIDELAKAICKKIKNDVSLDNEADHNGSLISGAYLLSNLTKKISAIKDSGEIK 292
DB 600 ASQSLGSLDNLNNAOKQOYTDQINGAHYVDEANQIKONQNLMTAGNKLQAATADKATK 659
QY 293 AEI-----EKAK-----CSEFTAKIKGHTDLG--KEGVTDNNAKALKITNNDK 337
DB 660 ATVNETDADQAKOAVNTAVTNAENIISKANGSNATQAEVEQAIKOYNNAKQALNGNANV 719
QY 338 TKGADELEKLFESVKNLSKAKKEMLTNSVYELTS 371
```

Db 720 QHAKDEATALINSSNDLNQAKQALQOYONATT 753

RESULT 9

US-10-072-851-13080
; Sequence 13080, Application US/10072851

GENERAL INFORMATION:

APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles

APPLICANT: Bussey, Howard

TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits C

TITLE OF INVENTION: Proliferation

FILE REFERENCE: ELITRA 028A

CURRENT APPLICATION NUMBER: US/10/072,851

CURRENT FILING DATE: 2002-02-08

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

NUMBER OF SEQ ID NOS: 15811

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 13080

LENGTH: 875

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-10-072-851-13080

Query Match 10.3%; Score 187; DB 7; Length 875;

Best Local Similarity 22.2%; Pred. No. 4.6e-05;

Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

Db 11 NTSANSADESVKGPLTPEISKRTDS-----NAVLLAVEKVEALLSIDELAKAIGKKIK 65
Db 315 NTAATNLKNGIQDQNTTIGQVNFDTADAKRVAATNAVTAQAEQIIL-----KAQGPNTS 368
Qy 66 NDG-----SLDNEANRNEISLAGAYTTISTITOKLSKNGS--EGLKEKIAAKKC 114
Db 369 KQGVETALENVQRAKNEINLNGN--VANAKTAKNALNLNLTSTNNQKELKQIGCATTV 427
115 S--EEFSTKLKD-NHAQLGIG--GVTDENAKKAILKANAAAGKDGVELEKLS-----G 163
Db 428 AGVNVQSTTASSELMTAMNGLNGINDEAATKAAL-----NGTQNELEKAKQAHNTAID 479
Qy 164 SLESISAKAKEMLVKVELTSPVYHGNNSGKDGNTSA-----NSADESV--KGPNT 214
Db 480 GLSHLTNNQKALQOLVQOSTTVAEAGNEQKANNVNDAMDRLKROSIADNATTKQNONYT 539
Qy 215 EIS--KRITESNAV-----LAVKE 232
Db 540 DASQNKKAATNAVYTAGIITDQTSPTLPYVINOAGQVSTTKNALNGENLEAKQ 599
Qy 233 VETLTSTIDELAKAIGKIKNDVSLDNEADHNGSLISGAYLISNLTKKISAIKDSGLK 292
Db 600 ASOSIGSIDNNMNAQKQVTVTDQINGAHVDEANQIKONNONTAMGNLQAIADKDATK 659
Qy 293 AEI-----EKAKK-----CSEEFKAKLGEHTDLG--KRGVYDDNAKKAILKTNNDK 337
Db 660 ATVNFTDADQAKQOAYNTAVTNAENIISKANGNATQAEVEQAIQVNAKALGNANV 719
Qy 338 TKGADELEKLEFESVKNLSKAKEMLTNSVKELTS 371
Db 720 QHAKDEATALINSSNDLNQAKQALQOYONATT 753

RESULT 10

US-10-032-585-7646
; Sequence 7646, Application US/10032585

GENERAL INFORMATION:

APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7646
LENGTH: 1881

TYPE: PRT

ORGANISM: Candida albicans

FEATURE:

NAME/KEY: MISC_FEATURE

LOCATION: (1881)..(1881)

OTHER INFORMATION: X-any amino acid

US-10-032-585-7646

Query Match 10.1%; Score 182.5; DB 7; Length 1881;

Best Local Similarity 22.6%; Pred. No. 0.00026;

Matches 101; Conservative 75; Mismatches 139; Indels 131; Gaps 20;

Qy 26 LPEISKRT---DSNAVLL-AKVEVALL-----SSIDELAKAIGKKIKND 67
Db 1443 LPEATSELTAKIDNNQSLTEIEKTKAALTKSSKDLVCGNOKSLQDSLSKSVSEKLN- 1501
Qy 68 GSLDNEANRNEISLAG-----AYTISTITOKLSKL-----NGSEGLKERTIAA 111
Db 1502 --FENKYNQETSLKDEIEKQKEIYTLQTELKDRISVEKERAAMLSSENSEYIKREYSK 1559
Qy 112 KCCSEEFSTKLKDNNHAQLGIGVTDENAKKAILKANAAAGKDGVELEKLSGLESLSKA 171
Db 1560 IKSLESKINSIKENHS-----KEITTHNEQKTSLK-----QDIAKLSQDHS-AQT 1604
Qy 172 AKEMLANSVKELTSPVYHGNNSGKDGNTSANSADSVKGPLTPEISKI-----TES 223
Db 1605 QLEDKNEQKELKASL-----EKHNTESATSIIE--KNNQIKLESETTKSLKTELKTSG 1656
Qy 224 NAVVLAKEVEVETLTSTIDELAKAIGK-----KIKNDVSLDNEADHNGSLISGAYLISNL 277
Db 1657 DALKOSQKEYTKLTKRNSDPTESKLEKQLEBELEKVKSDLOTADK-KLKGITEREIALKSEL 1715
Qy 278 ITRKISAIKDSGLKA-----EIERAK-----KCEEFKAKLGEHTDLG-- 318
Db 1716 ETVKNSGLTSTSELALTTKVSLEKEKELOFLSGNKSSELDYIQR-----HSDISEKL 1771
Qy 319 EGYTD-----DNAKAILKTNDKTKGADEL----- 345
Db 1772 KALTDLEKKEKQFDSDSKTLELNDLTSTKKELETEKTQTSKFNRLERKDKIEYKLN 1831
Qy 346 KLFESVKNLSKAKEMLTNSVKELTS 371
Db 1832 KELELLKNDSGAKKELSEKVKLES 1857

RESULT 11
US-10-072-851-15590
; Sequence 15590, Application US/10072851

GENERAL INFORMATION:

APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.

```

; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15590
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-072-851-15590

```

```

Query Match          10.1%; Score 182.5; DB 7; Length 1881;
Best Local Similarity 22.6%; Pred. No. 0.00026;
Matches 101; Conservative 75; Mismatches 139; Indels 131; Gaps 20;

```

```

QY 26 LTESKRTT---DSNAVL-AYKEVEALL-----SSIDELAKAIGKIKND 67
DB 1443 LTESLTKLQDNQSLTEIEIEKTKALTKSSKDEYCGNQKSLQDSLTKSVKSELKN- 1501
QY 68 GSLNEANRNESLAG-----AYTSTLTQKLSL-----NGSEGLKETAIAA 111
DB 1502 --FENKQOETSLKDEIEEKQKELVTLQTELKDRISIEVEKERAMLSSENSTVIKEYSDK 1559
QY 112 KCSSEFTKLKDNHAQIGVTDENAKAILKANAGKDKGVELEKLSGSLESLSKA 171
DB 1560 IKSLESKINSIKENHS---KEITTHNQKTSK-----QDLAKLSQDHES-AQT 1604
QY 172 AKEMLAISVKELTSPVYHGNNSGDKGNTSANSADSVKGNPLTETSKT-----TES 223
DB 1605 QLEKEMQLELKASL-----EKHNTESATISIE--KNQIKELSETIKSLKTELKTSG 1656
QY 224 NAVLAAYKEVETLTSIDELAKAIGK-----KIKNDVSLDNEADHNGSLISGAYLISNL 277
DB 1657 DALKQSOKEKTKLTKNSDYESKLEKQLEBEKVKSDLOTADE-KLKGITEREIALKSEL 1715
QY 278 ITKKLSAIKSGELKA-----EIERAK-----KCSSEFTAKLKGHTDLCK-- 318
DB 1716 EYKNSGLSTSELALTKTVKSLKLEKEBELQFLSGNKSKELEBYIQK---HSDISEKL 1771
QY 319 EGYVD-----DNKKAIIKTNNDKTKGADELE----- 345
DB 1772 KALTDELKERTKQFDDSKKLTLELNDLITKKELETKTQTSKFNKLEBKDEIVKLN 1831
QY 346 KLFESYKNIKSKAKEMLTNSVKELTS 371
DB 1832 KELELLKNDNSGAKKELSEKVSLES 1857

```

```

RESULT 12
US-09-708-427-19883
; Sequence 19883, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P

```

```

; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..1014
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc-feature
; LOCATION: 1..1014
; OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

```

```

Query Match          9.8%; Score 176.5; DB 6; Length 1014;
Best Local Similarity 24.4%; Pred. No. 0.0003;
Matches 93; Conservative 81; Mismatches 152; Indels 55; Gaps 16;

```

```

QY 16 SADESIV--KGNPLTETSKITDSNAV--LLAVKEVEALLSSIDEL-----AKAIGKKIK 65
DB 516 SAESLQKRGREIDAVTKRMELAHQSLSDSEHRLQKMEEFTRSDSASSLTKIR 575
QY 66 NDGSLDNEANRNESLAGVITSTLTQKLSKLSGSEGLKETAIAAKCSSEFTKLKN 125
DB 576 D---LEKINSYEQULAEASGSSSLKEKLEQTLG-----RLAASVNE---KIKOE 622
QY 126 HAQIGVTDENAKAILKANAGKDKGVELEKLSGSLESLSKAKEMLAISVKELTS 185
DB 623 FDQOQEKSLQSSSESLAETNNQKTK-IOELGLIGSSVEKETRAKLEAEATERFQ 681
QY 186 PYVHGNNSGDKGNTSANSADSVKGNPLTETSKITDSNAVVL--AYKEVETLTSIDEL 243
DB 682 KETESDVLKELTKHENQIEYK---LAHAGSGVADTRREVELEDAKSLKLNESTTEEL 738
QY 244 A-----KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKSIAIDSELSKA-EI 295
DB 739 GACQGLEKSSGLAEVNLNLNLELNNHS-----EAVELDTKLSALEAEKEQTANEL 791
QY 296 EKAKKCSSEFTAKL--KGEHTDLKGEVTDENAKAILKTNNDKTKGADELEKLFESVK- 352
DB 792 EAKTTEIDLTQTLTSGEKLOQIEKLRVAVAEKVLSH-----FEELEKTLSEVNA 845
QY 353 NLSKAKEMLTNSVK--ELTS 371
DB 846 QLEKVENNATASVKAELTS 866

```

```

RESULT 13
US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19882
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..1018
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc-feature
; LOCATION: 1..1018

```

OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19882

Query Match 9.8%; Score 176.5; DB 6; Length 1018;
Best Local Similarity 24.4%; Pred. No. 0.0003;
Matches 93; Conservative 81; Mismatches 152; Indels 55; Gaps 16;

QY 16 SADES- -KGNLITEISKTTDSNAV--LLAVKEVALLSIDEL-----AKAIGKKIK 65
DB 520 SAESLEOKGREIDEATTKMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSITEKLR 579
QY 66 NQSLDNEANRNESLAGAVTTISTLTOKLSKNGSEGLKEKIAAKKCEESTRTKDN 125
DB 580 D---LEGKIKSTEBQLAESGKSSLEKEQTGLG-----RLAAESVNE---KLKOE 626
QY 126 HAQLOGVTDENAKKAILKANAAGKDKGVEELEKLSLESLSKAKEMIANSEKELTS 185
DB 627 FDQAGKSLQSSSESESLAETNNQKIK-IQLEGLIGSGSVEKETALKRLEAIERFNQ 685
QY 186 PVYHGNSGKDGNTSANSADSVKGNLITEISKTTESNAVVL--AVKEVETLLTSIDEL 243
DB 686 KETESSDLVEKLTHTHENOIEYK---LAHESGVADTRKVELEDALSKLNLESTIEEL 742
QY 244 A-----KAIGKKIKNDVSLDNEADHNGSLISGAVLISLITKIKSAIKDSGELKA-EI 295
DB 743 GAKCQGLEKESGDLAEVNLKLNLELANHGS-----EANELQTKLSALEKEQTANEL 795
QY 296 EKAKCSEETFAKL--KGEHTDLKKEGVTDNNAKAILKTNNDKTKGADELKLFESVK- 352
DB 796 EAKSTIEDTLTKQTSBEGKLOQSIKRLAVAAKSVLESH-----FEELKLTSEVKA 849
QY 353 NLSKAKEMLTNSVK--ELTS 371
DB 850 QLKENVENATAVYKVAELTS 870

RESULT 14
US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1269
; OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881

Query Match 9.8%; Score 176.5; DB 6; Length 1269;
Best Local Similarity 24.4%; Pred. No. 0.0004;
Matches 93; Conservative 81; Mismatches 152; Indels 55; Gaps 16;
QY 16 SADES- -KGNLITEISKTTDSNAV--LLAVKEVALLSIDEL-----AKAIGKKIK 65
DB 771 SAESLEOKGREIDEATTKMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSITEKLR 830
QY 66 NQSLDNEANRNESLAGAVTTISTLTOKLSKNGSGLKEKIAAKKCEESTRTKDN 125

DB 831 D---LEGKIKSTEBQLAESGKSSLEKEQTGLG-----RLAAESVNE---KLKOE 877
QY 126 HAQLOGVTDENAKKAILKANAAGKDKGVEELEKLSLESLSKAKEMIANSEKELTS 185
DB 878 FDQAGKSLQSSSESESLAETNNQKIK-IQLEGLIGSGSVEKETALKRLEAIERFNQ 936
QY 186 PVYHGNSGKDGNTSANSADSVKGNLITEISKTTESNAVVL--AVKEVETLLTSIDEL 243
DB 937 KETESSDLVEKLTHTHENOIEYK---LAHESGVADTRKVELEDALSKLNLESTIEEL 993
QY 244 A-----KAIGKKIKNDVSLDNEADHNGSLISGAVLISLITKIKSAIKDSGELKA-EI 295
DB 994 GAKCQGLEKESGDLAEVNLKLNLELANHGS-----EANELQTKLSALEKEQTANEL 1046
QY 296 EKAKCSEETFAKL--KGEHTDLKKEGVTDNNAKAILKTNNDKTKGADELKLFESVK- 352
DB 1047 EAKSTIEDTLTKQTSBEGKLOQSIKRLAVAAKSVLESH-----FEELKLTSEVKA 1100
QY 353 NLSKAKEMLTNSVK--ELTS 371
DB 1101 QLKENVENATAVYKVAELTS 1121

RESULT 15
US-09-708-427-15046
; Sequence 15046, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15046
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1144
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1144
; OTHER INFORMATION: Ceres Seq. ID 1828629
US-09-708-427-15046

Query Match 9.4%; Score 169.5; DB 6; Length 1144;
Best Local Similarity 29.3%; Pred. No. 0.0011;
Matches 110; Conservative 42; Mismatches 154; Indels 69; Gaps 18;

QY 15 NSADESVKGNLITEISKTTDSNAVLLAVKEVALLSIDELAKAIGKKIKNDGSLDNEA 74
DB 606 NLVDNVANMONTIAEESDRLR-----REV-AYLKIDIDSLTANGGLADVTNLOIS 656
QY 75 NRNE-----SLAGAVTTISTL--ITOKLSK-----NGSEGLKEKIAAKKCEEST 120
DB 657 EENKELERETTLKKAELSLSLVDKASKIQTVQENBELRERETATLKRIEELSK 716
QY 121 K---LKDNHQLOGVTDENAKKAILKANAAGKDKGVEELEKLSLESLSKAKEMLA 177
DB 717 LHEILSQETRIQI-----SNHEKEELKRETAAYLKRIEELSKV-----QEDLL 760
QY 178 NSVKELTSPVYHGNSGKDGNTSANSADSVKGNLITEISKTTESNAVVL--LAVKE 232
DB 761 NKENELHGVVEIEDL-RSKDSLAKOKRIEELSNFNASLLIKE-NELDAVGCENELSKQ 818
QY 233 VETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAVL--ISNLTIKKISAIKDS 288
DB 819 VSTLTQ-IDELSDLKQSLIKHEKELQAAIYENETLKAEALSLORIELTTLTKQTLIDKQ 877

OY 289 GELKA-----EIEKAKCS--EFTAKLGEHTDCKEG----VDDNKAAILKTND- 336
DB 878 NELQGVFHEHELELAKAKASSIKRIDELLHLEQSWLEKESEFQVTOENLE--LKTQDAL 934
OY 337 KTRGADELEKLFESV 351
DB 935 AAKTIEELSKLKESTL 949

Search completed: March 18, 2002, 09:58:23
Job time: 381 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:47 ; Search time 68.77 Seconds

(without alignments)
410.946 Million cell updates/sec

Title: US-09-596-746a-34

Perfect score: 1809

Sequence: 1 MACNSGRKDGNTSANSADSE.....KNLSKAKEMLTNSVKELTS 371

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	884	48.9	193	2	S70280	outer surface prot
2	883	48.8	211	2	S69918	outer surface prot
3	879	48.6	210	2	I40144	outer surface prot
4	862	47.7	203	2	I40108	outer surface prot
5	849	46.9	178	2	I40106	outer surface prot
6	663.5	36.7	210	2	S69927	outer surface prot
7	662	36.6	193	2	S70287	outer surface prot
8	658.5	36.4	210	2	S70218	outer surface prot
9	656	36.3	200	2	I40122	outer surface prot
10	654	36.2	207	2	S69919	outer surface prot
11	654	36.2	193	2	S70286	outer surface prot
12	653	36.1	207	2	S37727	outer surface prot
13	651.5	36.0	193	2	S70276	outer surface prot
14	649.5	35.9	212	2	I40143	outer surface prot
15	647	35.8	207	2	I40271	outer surface prot
16	643	35.5	189	2	S70263	outer surface prot
17	640	35.4	191	2	I40153	outer surface prot
18	638	35.3	200	2	S54198	outer surface prot
19	636.5	35.2	211	2	I40268	outer surface prot
20	636	35.2	209	2	S69917	outer surface prot
21	635	35.1	207	2	I40276	outer surface prot
22	628	34.7	191	2	S70284	outer surface prot
23	622	34.4	211	2	I40145	outer surface prot
24	621	34.3	193	2	S70279	outer surface prot
25	615.5	34.0	194	2	S70277	outer surface prot
26	613	33.9	192	2	S70285	outer surface prot
27	612	33.8	193	2	S70259	outer surface prot
28	612	33.8	212	2	S69922	outer surface prot
29	610	33.7	176	2	I40121	outer surface prot

30	609.5	33.7	190	2	S70273	outer surface prot
31	608	33.6	210	2	S69923	outer surface prot
32	608	33.6	212	2	I40279	outer surface prot
33	607.5	33.6	177	2	I40129	outer surface prot
34	607	33.6	194	2	S70268	outer surface prot
35	606	33.5	178	2	I40125	outer surface prot
36	605.5	33.5	193	2	S70274	outer surface prot
37	603	33.3	178	2	I40104	outer surface prot
38	602.5	33.3	199	2	S54197	outer surface prot
39	599	33.1	209	2	I40273	outer surface prot
40	597	33.0	191	2	S70278	outer surface prot
41	595.5	32.9	211	2	S69932	outer surface prot
42	592	32.7	209	2	I40281	outer surface prot
43	590.5	32.6	214	2	S69916	outer surface prot
44	589.5	32.6	194	2	S70289	outer surface prot
45	588	32.5	209	2	I40142	outer surface prot

ALIGNMENTS

```
RESULT 1
S70280
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 12-Feb-1998 #sequence, revision 20-Feb-1998 #text, change 26-May-2000
C:Accession: S70280
R:Livley, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; PMID:96296448
A:Accession: S70280
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-193 <LIV>
A:Cross-references: EMBL:LA2868; NID:9858735; PIDD:AB37011.1; PTD:91695226
A:Experimental source: strain Z57
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 48.9%; Score 884; DB 2; Length 193;
Best Local Similarity 98.4%; Pred. No. 2,9e-34;
Matches 183; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 C NNSGRKDGNTSANSADSEVKGPNLTETSKRTTDSNAVLAVKEVEALSSIDELAKAIGK 62
|||||
Db 1 C NNSGRKDGNTSANSADSEVKGPNLTETSKRTTDSNAVLAVKEVEALSSIDELAKAIGK 60

Oy 63 KIKNDGSLDNEANRNESLAGAVTISTLTOKLSKLNGSEGLKRIIAAKKCEEFSTKL 122
|||||
Db 61 KIKNDGSLDNEANRNESLAGAVTISTLTOKLSKLNGSEGLKRIIAAKKCEEFSTKL 120

Oy 123 KDNNAOGLGIGVDENKAKKALKANAAGKDKGVLELTKLSLSLSKAAKEMLANSVKE 182
|||||
Db 121 KDNNAOGLGIGVDENKAKKALKANAAGKDKGVLELTKLSLSLSKAAKEMLANSVKE 180

Oy 183 LTSPVY 188
|||||
Db 181 LTSPVY 186

RESULT 2
S69918
outer surface protein C precursor - Lyme disease spirochete (strain PBre)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
A:Variety: strain PBre
C>Date: 06-Dec-1996 #sequence, revision 14-Feb-1997 #text, change 26-May-2000
C:Accession: S69918; S72674; I40103
R:Jaursis-Reipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
```

OY	191	NNSGDDGNMSAADESVKGNLPTEISKKITTESNAVVLAVKEVEETLLTSIDELAKAIGKR	250
Dd	20	NNSGGDGNMSAADESVKGNLPTEISKKITTESNAVVLAWEVEETLLTSIDELAKAIGKR	79
OY	251	IKNDVSLDNEADHNGSLISGAYLISNLTTRKISAINDSGELKAETIEKARKCSEEFARKL	310
Dd	80	IKNDVSLDNEADHNGSLISGAYLISLTTRKISAIKDSGELKAETIEKARKCSEEFARKL	139
OY	311	GEHDLDGEGSYTDNNAKAILTKTNNDKTKGADELKLEFESVKNLSSRAAKEMLTNSVKELT	370
Dd	140	GEHDLDGEGSYTDNNAKAILTKTNNDKTKGADELKLEFESVKNLSSRAAKEMLTNSVKELT	199
OY	371 S	371 S	
	I	I	
Dd	200 S	200 S	

RESULT 4
140108
Outer surface protein C - Lyme disease spirochete (strain DK7) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: 140108. S34176
R:Pfeisen, M.; Frederiksen, B.; Lebecq, A.M.; Vuust, J.; Hansen, K.
J. Clin. Microbiol. 31, 2570-2576, 1993
A>Title: Polymorphism in ospC gene of Borrelia burgdorferi and Immunoreactivity of Os
A:Reference number: 140105; MUID:94075528
A:Accession: 140108
A>Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-203 <RES>
A:Cross-References: EMBL:X73625; NID:g313375; PIDN:CAA52004.1; PID:g313276.
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Best Local Similarity 95.7%; Pred. No. 3, 1e-33; Length 205;
Matches 179; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
OY 1 MACNNSGKDGNTSANSADSESYKCPNLTEISKLTIDNSAVLLAKVEEALLSSIDELAKAI 60
Db 17 TSCNNSGKDGNTSANSADSESYKCPNLTEISKLTIDNSAVLLAKVEEALLSSIDELAKAI 76
OY 61 GKTIKNGDSLNDNEANRNESLIAGAYTITSLITOKLSKLGSEGLKXIIAAKRCSEEST 120
Db 77 GKTIKNGDSLNDNEANRNESLIAGAYTITSLITOKLSKLGSEGLKXIIAAKRCSEEST 136
OY 121 KTKNHQOLIGQYTDENNAKKALKANAAKDKQVDELEKLSGSLETSKAKEMLVANSV 180
Db 137 KTKNHQOLIGQYTDENNAKKALKANAAKDKQVDELEKLSGSLETSKAKEMLVANSV 196
OY 181 KELTSVPV 187
Db 197 KSLDSYV 203
```

RESULT 5
I40106

outer surface protein C - Lyme disease spirochete (strain H09) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40106; S54185
R:Pielsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelson, B.; Lebech, A.M.; Hansen, K.
J.; Bacteriol. 177, 3036-3044, 1995

A:Accession: I40106
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-178 <RES>

A:Cross-references: EMBL:X84765; NID:g793789; PIDD:CAA59236.1; PID:g793790
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match	46.98%	Score 849	DB 2	Length 178
Best Local Similarity	99.48%	Pred. No.	1e-32	
Matches 177; Conservative	1	Mismatches	0	Gaps 0

Qy	11	NTSANSADSEVYGCNLTLEISKRTIDNSAANVLLAVEVALLS1DELAKAIGKKIKNGSL	70
Db	1	NTSANSADSEVYGCNLTLEISKRTIDNSAANVLLAVEVALLS1DELAKAIGKKIKNDGSL	60
Qy	71	DNEANRNESLLAGAVTTSTLTQTKSLKNGSEGLKEKTAIAAKKCSEEFSTRLKDNDHAOLG	130
Db	61	DNEANRNESLLAGAVTTSTLTQTKSLKNGSEGLKEKTAIAAKKCSEEFSTRLKDNDHAOLG	120
Qy	131	IQGVTDENAKKAILLEANAAGKDKQVEBELKSLGSLJESLSKRAKEMLANSYKELTSPVY	188
Db	121	IQGVTDENAKKAILLEANAAGKDKQVEBELKSLGSLJESLSKRAKEMLANSYKELTSPVY	178

Result 6
S69927
outer surface protein C precursor - Lyme disease spirochete (strain Pk)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

Query Match	Score 663.5	DB 2	Length 210
Best Local Similarity	77.28	Pred. No. 3.8e-24	
Matches 146	Conservative 10	Mismatches 30	Indels 3
			Gaps 3

Qy	1	MACNNNSGKDGTSMANSADSESVKSGNLTLEISKRITDSNAVLTAAYKVEYALLSIDEL- AKa	59
Qy	1	MACNNNSGKDGTSMANSADSESVKSGNLTLEISKRITDSNAVLTAAYKVEYALLSIDEL- AKa	59
Db	17	ISCHNNSGKDGTSMANSADSESVKSGPNLTLEISKRTIDSNAAVLLAAYKVEYALLSIDEI	NAKa 76
Qy	60	IGKKRTIKNDSDLDNEANRNESILLAGAYITSTLTOKUSLNGSEGLKERTIAAAXKCEEF	S 119
Db	77	IGKKRIHQNNGDPTDENHNGSLLAGAYALSTLTIKOKLGOLK- NEGLKERTIEAAXKCE	ET 135
Qy	120	TKLKDHAQLOIGVYTDENAKKALLKANAAGKDKGVLELEKLSLSSLSKAAREMLANS	179
Db	136	NKLDKRHTDLKEGVTADAKAEALIKTNGT- KTRGAEBELGLTFESVEVLSKAAREMLANS	194
Qy	180	VKELTSPVY 188	
Db	195	VKELTSPVY 203	

RESULT 7
S70287
Outer surface protein C - Lyme disease spirochete
C.Species: Borrelia burgdorferi (Lyme disease spirochete)
C.Date: 12-Feb-1998 #sequence_rev1sion 20-Feb-1998 #text_change 26-May-2000
C.Accession: S70287
R.Livley, I.: Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A.Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A.Reference number: S70255; MUID:96296448
A.Accession: S70287
A.Status: nucleic acid sequence not shown
A.Molecule type: DNA
A.Residues: 1-193 #GLV>
A.Cross-references: EMBL:142895; NID:9858723; PIDD:AA837003.1; PID:91695220
A.Experimental source: strain 28354
C.Genetics:
A.Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match	36.6%	Score	662	DB 2:	Length	193:			
Best Local Similarity	75.9%	Pred	No	4.1e-24:					
Matches	142:	Conservative	17:	Mismatches	26:	Indels	2:	Gaps	2

0y	3	CNNSKCDTSTANSADSEVKNPNTLEISKRTDSTNAVLAVKEVAALSSIDEJLA-PAIG	61
Db	1	CNNSKCDTSTANSADSEVKNPNTLEISKRTSTESNAVLAVKELETLLASIDELATNAIG	60
0y	62	KKIKNDGSLDEANRNESLLAGAATTISTLTQKLSKLNGSEGLKEKIAAAKKCEEFSTK	121
Db	61	KKIQNGGLAVEAGHNGTLTLAGAATTISKLTQKIDGLTENSEKLEKTEIENAKCSEDEFTKK	120
0y	122	LKDHNAGIGCYVDENAKKAILANANGGRKGYBELEKLSSTLESJSAKKEMLANSVK	181
Db	121	LEGHAAGIGIENVYDENAKKAILITDAA-KDGAAELEKLEKRAVENLAKAKKEMLANSVK	179
0y	182	ELTSPVY 188	
Db	180	ELTSPIV 186	

RESULT 8

outer surface protein C - Lyme disease spirochete
670218

C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C:Accession: 670218; 140269; 537726; 570281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Krelavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
A:Reference number: A70100; MUID:98065943
A:Accession: 670218

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-210 <KLE>

A:Cross-references: GB:AE000792; NID:q3253098; PIDN:MAC66329.1; PID:q2689901; TIGR:BB

A:Experimental source: strain B31

R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995

A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu

A:Reference number: 140269; MUID:96025162

A:Accession: 140269

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-210 <RES>

A:Cross-references: GB:DA9497; NID:g707092; PIDN:BAA08457.1; PID:g769684

R:Jaunis-Heipke, S.; Fuchs, R.; Metz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Mol. Microbiol. Immunol. 182, 37-50, 1993

A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (osp

```
OY      121 K LKDNHQAOLGIQGVNDENAKRAIIILKANAAAGKDKGVEELKEJLGSLESLSKAEMKLANSV   180  
       |||||:-||||-|-:|||||::||| |::||| |::||| |::||| |::||| |::||| |::|||  
Db     134 KLKDSHAELGIGSVODDNNKKRATIKTHGT -KDKGAKELEEFLFKSLESLSKAQAALNTNSV   192
```

```
OY      181 KELTSPVV 188  
       ||||:-|||  
Db     193 KELTNPVV 200
```

```
RESULT    10  
569919  
outer surface protein C precursor - Borrelia garinii (strain P7rob)  
C:Species: Borrella garinii  
A:Variety: strain P7rob  
C>Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_Change 26-May-2000  
C:Accession: S69919  
R:Jarvis-Helpke, S.; Liegl, G.; Preac-Murusic, V.; Roessler, D.; Schwab, E.; Soutschek,  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A>Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre  
A:Reference number: I40047; MUID:95395018  
A:Accession: S69919  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-207 <>(JAN>  
A:Cross-references: EMBL:X83554; NID:g872027; PIDN:CA58544.L; PID:g872028  
A:Experimental source: strain P7ob  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995  
C:Genetics:  
A:Gene: ospc  
C:Superfamily: Lyme disease spirochete surface protein C  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-207/Product: outer surface protein C #status predicted <MAT>
```

```
Query Match          36.3%; Score 656; DB 2; Length 207;  
Best Local Similarity 75.0%; Pred. No. 8.3e-24;  
Matches 141; Conservative % 20; Mismatches 23; Indels 4; Gaps 3;
```

```
OY      1 MACNNGSGDGMTSANSDAESVKGPMLTFISKRITDSNAVLAVAEVAEALLSIDELAIAI 60  
       :::|||:-||::|||:-||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||  
Db     17 ISCNSS-<-GDASSTNPDESAKGPMLTVISKRTIDSNFLLAVEVELLISSIDELEKAI 74  
       ::|||:-||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||  
OY      61 GKRIKINDSLNEANKRNRESLAGAYTTISTLTIOKLSKNSSEGGEKRIAARKCESEFT 120  
       |||||:-||||:-||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||  
Db     75 GKRIKNOSTLDNEANKRNRESLAGAYEISKITLOKLSLVIN-SSELKEKIKEAKDCSEKFTT 133  
       |||||:-||||:-||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||  
OY      121 K LKDNHQAOLGIQGYVDENAKRAIILEKANAAAGKDKGVEELEKLGSLLESLSKAEMLANSV 180  
       |||||:-||||:-||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||  
Db     134 KLKDSHAELGIGSVODDNNKKRATIKTHGT-KDKGAKELEEFLFKSLESLSKAQAALNTNSV 192
```

```
OY      181 KELTSPVV 188  
       ||||:-|||  
Db     193 KELTNPVV 200
```

```
RESULT    11  
S70286  
outer surface protein C - Lyme disease Spirochete  
C:Species: Borrellia burgdorferi (lyme disease Spirochete)  
C:date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000  
C:Accession: S70286  
R:Lively, J.; Gibbs, C.F.; Schuster, R.; Dorner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:title: Evidence for lateral transfer and recombination in OspC variation in lyme di  
A:Reference number: S70255; MUID:96296448  
A:Accession: S70286  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-193 <LTY>  
A:Cross-references: EMBL:I42897; NID:g858728; PIDN:AAB37006.1; PID:g1695222  
A:Experimental source: strain 26B15  
C:Genetics:
```


RESULT 15

I40271

outer surface protein C precursor - *Borrelia garinii*

C;Species: *Borrelia garinii*

C:\Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000

C; Accession: I40271

R; Fukunaga, M.; Hamase, A

J. Clin. Microbiol. 33, 2415-2420, 1995

A;Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi* sensu lato
Reference number: TA0260, MIMD:06025163

A;Reference number: I40269; MUID:96025162

A;Accession: I40271

A;Status: preliminary; translated from GB/EMBL/DBJ
 1: Molecule type: DNA

A;Molecule type: DNA

A;Residues: 1-207 <RES>
A:Cross-references: CB:

A; Cross-references: GB:D49377; NID:g1041103; PIDN:BA08375.1; PID:g1041104
C: Superfamily: Lyme disease spirochete surface protein C

C; Superfamily: Lyme disease spirochete surface protein C

Query Match	Score	DB 2;	Length
35.88;	647;	DB 2;	207;
73.38;	647;	DB 2;	207;

Best Local Similarity 72.38; Pred. No. 2.1e-23/
Matches 136. Conservative 37. Mismatches 31

Matches	136;	Conservative	27;	Mismatches	21;	Indels	4;	Gaps	3;
---------	------	--------------	-----	------------	-----	--------	----	------	----

1 MACNSGKDNTSANSADSVKGPNLTEISKKITDSNAVLAVKEVEALLSSIDELAKAI 600
0Y

Db 17 ISCNNG--GDTASTNPDESAGPNLIEISKKITDSNAVVLAVKEVEALLSSIDELSKAI 74

QY 61 GKIKNDGSLDNEANRNESLAGAYTISTLLITQKLSKNGSEGLKEKIAAKKCSEEFST 120

Db 75 GKIRNDGTLNENRNESLIAGAYEISKLTÖKLSVLN-SEELKEKIKEAKDCSEKFTT 133

121 K L K D N H A Q L G I Q G V T D E N A K K A I L K A N A G K D K G V E E L E K L S G S L E S L S K A K E M L A N S V 180

Db 134 KL RDSHAELGVQNVQDDNAKRAILKTH-GNKDKGAKELKELSESLLENLAKAAQAASSNSV 1922

QY 181 KELTSPV 188

Db 193 KELTSPV 200

Search completed: March 18, 2002, 09:56:48
Job time: 371 sec

Job time: 371 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:49 ; Search time 39.62 Seconds
(without alignments)
343.328 Million cell updates/sec

Title: US-09-596-746a-34
Perfect score: 1809
Sequence: 1 MACNNSGKDNTSANSADSE.....KNLSRAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwlssProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	658.5	36.4	210	1	OSCL_BORBU
2	556	30.7	212	1	OSCL_BORBU
3	392.5	21.7	214	1	VM24_BORHE
4	373	20.6	215	1	VM03_BORHE
5	181	10.0	1957	1	YD86_SCHPO
6	180.5	10.0	1251	1	RBP2_PLAVB
7	178.5	9.9	1790	1	USOL_YEAST
8	173	9.6	564	1	M12_STRPY
9	171.5	9.5	1509	1	MYSN_ACACA
10	169.5	9.4	775	1	YHGE_BACSU
11	169.5	9.4	1947	1	MYSN_CAEEL
12	168	9.3	1937	1	MYH8_HUMAN
13	167	9.2	1966	1	MYSB_CAEEL
14	166.5	9.2	1713	1	LMAS_HUMAN
15	165.5	9.1	2869	1	RBP1_PLAVB
16	165	9.1	998	1	SCA4_RICAK
17	164.5	9.1	296	1	OSBI_BORBU
18	164.5	9.1	1938	1	MYH4_RABIT
19	163	9.0	1427	1	REST_HUMAN
20	163	9.0	1940	1	MYH3_CHICK
21	162.5	9.0	483	1	M6_STRPY
22	162.5	9.0	1433	1	REST_CHICK
23	162.5	9.0	3210	1	CENF_HUMAN
24	161.5	8.9	1526	1	MYS2_SCHPO
25	161.5	8.9	1941	1	MYH2_HUMAN
26	161	8.9	981	1	SCA4_RICEF
27	161	8.9	1084	1	MYSS_RABIT
28	160.5	8.9	1938	1	MYSS_RABIT
29	160	8.8	679	1	YKRP_YEAST
30	159	8.8	1934	1	MYH7_MESAU
31	159	8.8	1935	1	MYH7_HUMAN
32	157.5	8.7	1939	1	MYH1_HUMAN
33	157	8.7	1186	1	SMC_BACSU

34	157	8.7	1935	1	MYH7_PIG	P79293 sus scrofa
35	155.5	8.6	2349	1	TPR_HUMAN	P12270 homo sapien
36	155	8.6	1102	1	MISC_CHICK	P24616 gallus galli
37	154.5	8.5	465	1	MYH8_RABIT	P09323 drosophila
38	154.5	8.5	2017	1	MYSN_DROME	P08799 dictyostell
39	154.5	8.5	2116	1	MYS2_DICDI	0941t1 helicobacte
40	153.5	8.5	1167	1	CAGA_HELPJ	P34562 caenorhabdi
41	153	8.5	705	1	YNP9_CAEEL	0941t1 helicobacte
42	153	8.5	1013	1	SCA4_RICAK	P02564 ratius norv
43	153	8.5	1935	1	MYH7_RAT	P39922 hydra atten
44	151.5	8.4	539	1	MYS3_HYDAT	094623 homo sapien
45	151.5	8.4	1939	1	MYH4_HUMAN	

ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD:	PRT:	210 AA.
1	OSCL_BORBU	007337			
AC	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB19.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OG	Plasmid lp54.				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID:139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-93238332; PubMed-8478108;				
RA	Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an				
RT	major outer surface protein of Borrelia burgdorferi."				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-93238332; PubMed-8478108;				
RA	Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an				
RT	major outer surface protein of Borrelia burgdorferi."				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-96025162; PubMed-7494039;				
RA	Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;				
RA	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi."				
RL	Infect. Immun. 61:5097-5105(1993).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-96025162; PubMed-7494039;				
RA	Fukunaga M., Hamase A.;				
RA	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan."				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE-96025162; PubMed-7494039;				
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Ishihara R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kellavag A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,				

RX STAIN-SSP. H51 SEROTYPE 24;
 RA MEDLINE-9313310: Pubmed-1484486;
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
 RT "Subtelomeric expression regions of *Borrelia hermsli* linear plasmids
 RT are highly polymorphic."
 RL Mol. Microbiol. 6:329-331(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP3.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/aib-sib.ch>).
 CC or send an email to license@isb-sib.ch).
 CC -----
 DB EMBL: L04786; AAA22964.1; "-
 DB InterPro: IPR001800; Lipoprotein_6.
 DB Pfam: PF01441; Lipoprotein_6; 1.
 DB ProDom: PD001145; Lipoprotein_6; 1.
 DB DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DB Outer membrane; Lipoprotein; Signal; Plasmid.
 FT SIGNAL 1
 FT CHAIN 19 214
 FT VARIABLE MAJOR OUTER MEMBRANE
 FT LIPOPROTEIN 24
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 214 AA: 22541 MW: 415835F5102467C7 CRC64:

Query Match	21.7%	Score 392.5;	DB 1,	Length 214;
Best Local Similarity	47.6%	Pred. No. 2.4e-11;		
Matches 90;	Conservative 32;	Mismatches 54;	Indels 13;	Gaps 3;

0y	191	NSNGKDWTSANSADESVCBP----	NLTETSKITTESNAVLAKEXVETLTSTIDELAKA	246
Db	20	NNGGE-----LKSEYAKSDGTVLIDLAKSKKIKESAPASVCYKETTLYKSYDELAKA		74
0y	247	IGKKIKNDVSLDNADHNGLSICAVLYISNLTFKKIAIKD----	SGELAEITEKAKCS	302
Db	75	IGKKIKNDGGLDTEAGONGSLIAGVHSVSAVKTIVALETTSGISINELTKTKITEVSKA		134
0y	303	EEFPAKLGKEHTDLCGEGVTDNNKKAHLKTNNDKTKGADBLETKLFPSVKNLSKAAKEML		365
Db	135	EAPLAKLKDGHTLEGGKDKASDDTKKAIKDKNSDKTGAISELEKLNTPVADLLKPAAGVEV		194

363 TNSVKELTS 371
:::111:
195 EAIKELTA 203

	RESULT	4
WM03_BORHE		
ID	WM03_BORHE	STANDARD; PRF; 215 AA.
AC	002448;	
DT	01-JUL-1993 (Rel. 26, Created)	
DT	01-JUL-1993 (Rel. 26, Last sequence update)	
DT	01-FEB-1994 (Rel. 28, Last annotation update)	
DE	VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.	
GN	WMP3.	
OS	Borrelia hermsli.	
OC	plasmid.	
CC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.	
OX	NCBI_TaxID=140;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SSP. HSI SEROTYPE 3;	
RX	MEDLINE=93133110; PubMed=184486;	
RA	Restrepo B.T., Kitten T., Carter C.U., Infante D., Barbour A.G.;	
RT	"Subtelomeric expression regions of Borrelia hermsli linear plasmids	

```

FT      are highly polymorphic."
RL      Mol. Microbiol. 6:3299-3311(1992).
CC      -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC      FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC      -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC      ANCHOR.
CC      -1- SIMILARITY: STRONG. TO YMH24.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      The European Bioinformatics Institute. There are no restrictions on ways
CC      CC use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/)
CC      or send an email to license@isb-sib.ch.
CC
DR      EMBL; L04789; AAA22967.1; -.
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
DR      PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.
KW      Outer membrane; Lipoprotein; Signal; Plasmid.
FT      SIGNAL
FT      CHAIN
FT      1          18
FT      19         215
FT      VARIABLE MAJOR OUTER MEMBRANE
FT      LIPOPROTEIN 3.
FT      LIPID
FT      19         19
FT      SEQUENCE   215 AA;  23139 MW;  684C7AD35F87C771 CAC64;
N-ACYL DIGLYCERIDE (PROBABLE).

```

Query Match	20.68;	Score 373;	DB 1;	Length 215;
Best Local Similarity	46.38;	Pred. No. 1.7e-10;		
Matches	88;	Conservative 32;	Mismatches 56;	Indels 14;
			Gaps	4

[illegible]

RESULT 5
YD86_SCHPO

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYDROTHERMAL 222.8 KDA PROTEIN CLF3.06C IN CHROMOSOME I.
GN SPAC1F3.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use.


```

RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick R.E.;
RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEATITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X54378; CAA38253.1; -
DR EMBL: L03188; AAB00143.1; -
DR EMBL: U53668; AAB66659.1; -
DR PIR: A38455; A38455.
DR HSSP: P80220; 1DIP.
DR SGD: S0002216; USO1.
DR InterPro: IPR002017; Spectrin.
KM Transport: Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT DOMAIN 1 724
FT DOMAIN 725 1790
FT DOMAIN 465 487
FT DOMAIN 991 1790
FT DOMAIN 1172 1786
FT CONFLICT 847 847
FT CONFLICT 924 924
FT CONFLICT 1253 1253
FT CONFLICT 1319 1319
FT CONFLICT 1461 1461
FT CONFLICT 1581 1581
FT CONFLICT 1600 1600
FT CONFLICT 1661 1661
FT CONFLICT 1772 1772
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 9.9%; Score 178.5; DB 1; Length 1790;
Best Local Similarity 23.7%; Pred. No. 0.54;
Matches 89; Conservative 68; Mismatches 156; Indels 63; Gaps 13;

```

```

OY 307 AKIK-----GERTDGLKGEVTDNNAKKAIKTNNDKTKADE-LEKL-----PESVK 352
DB 1310 TELKIQLEKITNLNLSKAKKESKESLSRLKTSSEKRNKAEQLEKNEIQKNOAFKER 1369
OY 353 NLSKAKKEMLTNSVKE 368
DB 1370 KLNKSGSSTTQEXSE 1385

RESULT 8
M12_STRPY STANDARD; PRT; 564 AA.
AC P19401;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).
GN EMM12.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CS24 / Serotype M12;
RX MEDLINE=88058777; PubMed=2445730;
RA Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;
RT "Streptococcus pyogenes type 12 M protein gene regulation by upstream
RT sequences."
RL J. Bacteriol. 169:5633-5640(1987).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC PHAGOCYTOSIS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -1- SIMILARITY: TO OTHER M PROTEINS.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M18269; AAB88573.1; -
DR PIR: A60115; A60115.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003345; M_repeat.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02370; M; 9.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KM Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW Transmembrane; Coiled coil; Signal.
FT SIGNAL 1 41
FT CHAIN 42 564
FT DOMAIN 42 550
FT TRANSMEM 551 564
FT DOMAIN 44 505
FT DOMAIN 505 541
FT DOMAIN 542 547
FT NON_TER 564 564
SQ SEQUENCE 564 AA; 62904 MW; 5F1549DCA7B46 CRC64;

Query Match 9.6%; Score 173; DB 1; Length 564;
Best Local Similarity 23.1%; Pred. No. 0.26;
Matches 100; Conservative 56; Mismatches 149; Indels 128; Gaps 19;

```

QY 10 GNTSANSADSVKGNLITTEISKITDSNAVLAVKEVEALLSSIDELAKAI--GKKIKND 67
 DB 115 GRTGINDADLAK--ITTELEKSVSEKNOVLSQIKK-----ELEAEKDIOGREYH-- 163
 QY 68 GSDINDANRRESLACAYITITITK--LSKLINGS--EGLKEKIAAKKCSFEFTKUK 123
 DB 164 -AADLLRHQOE-----IAKEENVISKLNGELOPLKQKVDETRNLQOEKQVYL 210
 QY 124 DNHAQIGIOGVENDENAKKAILKANAGKDKGVLEELKSGSLESLSKAKEMLANSVKL 183
 DB 211 SLEQOL--AVYENKAKPELA-ALGHOLADREYNAKIAELSES-----KLADAKKDF 259
 QY 184 TSEPVV--HGNV-----SGKDS-----NTSANSADSVKGP----- 211
 DB 260 ELAALGHQAHNREYQAKLEKDKQIQIKQLEQOIIDASKGARADLEAVRAKKAETEL 319
 QY 212 -NLTEISKITDSNAVLAVK-----EVEITLTSIDE--LAKAIGKKI 251
 DB 320 NNLKAEIARVTEOKQILDSRKGTARDLEAVRKSQKQVEALKOLEQNKISEASRKGL 379
 QY 252 KNDVSLDNADHNGSLISGAYILSNLITKKISAIKDSGSELKAEIKAKCSEFTAKLKG 311
 DB 380 RRDLOTSRA-----KQOVEDLANLITELDKVKEKOISDASRQG 420
 QY 312 EHTDLGKEGVTDNNAKAILKTNNDKTCGADELKELFESVKNL-----SK 356
 DB 421 LRRLDASREAKQYKEA-LEBANSKILAEKLNKDLSESKLITEKAKELQAKLEAEK 479
 QY 357 AAKEMITNSVKEL 369
 DB 480 ALKEQALAKQAEEL 492

RESULT 9

ID MYGN ACACA STANDARD: PRT: 1509 AA.

AC P05659.
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
 OS Acanthamoeba castellanii (Ameoba).
 NCBI_TaxID=5755;

RP SEQUENCE FROM N.A.

MEDLINE=87308395; PubMed=3040773;

RT Hammer J.A., III, Bowers B., Paterson B.M., Korn E.D.;

RT "Complete nucleotide sequence and deduced polypeptide sequence of a
 RT hinge in the rodlike tail."

RL J. Cell Biol. 105:913-925(1987).

CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN

CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (MLC-2).

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
 CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE

CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.

CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
 CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMIC ACTIVITY.

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: Y00624; CAA68663.1; .
 DR PIR: A27224; A27224.
 DR HSSP: P08799; IAMD.
 DR InterPro: IPRO00048; IO.
 DR InterPro: IPRO01609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS00096; IO; 1.
 KW Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
 KW Methylation; Alkylation; Phosphorylation; Multigene family.
 FT DOMAIN 1 789
 FT DOMAIN 848 1509
 FT DOMAIN 848 1226
 FT DOMAIN 1227 1252
 FT DOMAIN 1253 1509
 FT DOMAIN 1253 1482
 FT DOMAIN 1483 1509
 FT NP_BIND 182 189
 FT DOMAIN 660 682
 FT DOMAIN 766 780
 FT MOD_RES 133 133
 FT MOD_RES 700 700
 FT MOD_RES 1489 1489
 FT MOD_RES 1494 1494
 FT MOD_RES 1499 1499
 SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE5173017E CRC64;

Query Match

Best Local Similarity 9.5%; Score 171.5; DB 1; Length 1509;
 Matches 97; Conservative 70; Mismatches 143; Indels 89; Gaps 18;

QY 19 ESVKGNPILTEISKITDSNAVLAVKEVA---LLSSIDELAKAIGKKIKNDGSLDNEA 74
 DB 953 EAERGELKASLEEEERNRKALQOEAKTVESERNELQOKEYEDEAAHDSLKKEEDLSREL 1012
 QY 75 NRNSILAGATTITLTLOKLSL--NSEGLK---EKIAAKKCSFEFTKLDNNAQLG 130
 DB 1013 RETDALADENISSETILRSKIKNTERGADVNRNLDVTAFTKLOLETKKSLSEELAQTR 1072
 QY 131 IOGVTDENAKKAILKANAGKDKGVEELE-----KLSGLESLSKAKEMLAN 179
 DB 1073 AQLEEKSGKEA---ASSKAKQLG-QQLEPARSEVDSLKLSNAEKSL-KTAKDONRDL 1127
 QY 180 VKELTSPVYHGNNSGKGNFSANSADSVKGP--PNLITEISKITT---ESNAVLAVKEV 233
 DB 1128 DEOL-----EDERTVRANVDQKQKALBAKLELDELDVATLDGQKNA---AAQA 1173
 QY 234 ETLTSTIDEL-----AKAIGKKIKNDVSLDNADHNGSLISGAYILSNLITKIS 283
 DB 1174 KILKQVDETKRRLAEASAAERLEKERN--ALDEVAQ-----LTAIDL 1216
 QY 284 AAKDSG-----ELKAEIERAKK---CSEFTAKLGEHTDLAGEGVTDNNAK 328
 DB 1217 AERDSGAQQRKRLNTRISLISQSELENAPKTGASSEEVKRLBEELRLBELITLAQEAAR 1276
 QY 329 AILKTNNDKTKGADLEKLFESYKNSKAKEMLTVSVK 367
 DB 1277 AAEK-NIDKAN--LELEELRQEADDAARDNDKLVKNRK 1312

RESULT 10
 ID YHGE_BACSU STANDARD: PRT: 775 AA.
 AC P32399;

DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00242; MYSC; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 855 MYOSIN HEAD-LIKE.
 FT DOMAIN 856 1947 ROD-LIKE TAIL (S2 AND LMM DOMAINS).
 FT NP_BIND 856 1947 COILED COIL (POTENTIAL).
 FT DOMAIN 174 181 ATP (BY SIMILARITY).
 FT DOMAIN 668 690 ACTIN-BINDING.
 FT MOD_RES 125 135 METHYLATION (TR1-) (POTENTIAL).
 FT MOD_RES 708 718 ALKYLATION (SH-1).
 FT MOD_RES 718 718 ALKYLATION (SH-2).
 FT CONFLICT 132 132 E -> D (IN REF. 2).
 FT CONFLICT 137 137 M -> I (IN REF. 2).
 SEQUENCE 1947 AA; 223009 MW; 6D5AE99BD013627 CRC64;

Query Match 9.4%; Score 169.5; DB 1; Length 1947;
 Best Local Similarity 22.9%; Pred. No. 1.5; Indels 135; Gaps 18;
 Matches 93; Conservative 61; Mismatches 118;

QY 23 GPNLEISKRTDSNAVLLAVKEVEALLSSIDELAKAIGKIKKNDG-----SID 71
 DB 1202 GQDLTVLKKKGSDA-----IQLSPDQIEQLQKQRIKKGKMGHOREDESCAALD 1252
 QY 72 NNA--NRNSLLAGATYST-----LITQKLSKLGNS-----EGIK 105
 DB 1253 QAKRLADDERLAKGEVYTSERLRKADQSHOLDPVSCK--GRNSENSDLARVEEIE 1311
 QY 106 EKIAAKKCESEFSTRKKNHQAOLGIVTDENAKKAILKANAAAGKGVLEKLGSL 165
 DB 1312 AKIQANRLKLOFNSML--DHAKRAEESRRKQNLNLSKILA-----RELEQKESI 1363
 QY 166 F-----SLSKAKEMLANSVKELTSPVHGNSGKDNTSANSADSVKGP-- 212
 DB 1364 EDEVAGKNENASHOLSKASVELDQWRTKFT-----EGLIGADPEDEKVRQK 1412
 QY 213 LFEIKKITESAVVLAKEVEITLTSIDELAKAIGKIKKNDVSDNADNHSLSGAY 272
 DB 1413 TSEIDALDACNAKIVALEMARSRLT-----AEADANRL-----EAHHQAQVSS-- 1457
 QY 273 LISNLTFRKISAI-----KDSGELKAIEKAKKCESEFTA---KKEGHTDGGKGV 321
 DB 1458 -----LEKKQKAFKVIDEMKKKVDVILELDGAQRDARQLSGEAKHLRGCHDTL----- 1507
 QY 322 TDDNAKKAAILKTNDKTGADLEKLFESVKNLSKAKEMLTNSVKE 368
 DB 1508 -----ADQVEGLRREKNKSLSDETRD-LTESISE 1534

RESULT 12
 MYH8_HUMAN STANDARD; PRT; 1937 AA.
 AC P13535; Q14910;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MYOSIN HEAVY CHAIN, SKELETAL MUSCLE, PERINATAL (MYHC-PERINATAL).
 GN MYH8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=90323631; PubMed=2373371;
 RA Karsch-Mizruchi I., Peghail R., Shows T.B. Jr., Leinwand L.A.;
 RT "Generation of a full-length human perinatal myosin heavy-chain-

RT encoding cDNA.";
 RL Gene 89:289-294(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=95324556; PubMed=7601129;
 RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schaffino S.,
 RA Steadman H.H., Rubinstein N.A.;
 RT "Characterization of a human perinatal myosin heavy-chain
 transcript.";
 RL Eur. J. Biochem. 230:1001-1006(1995).
 RN [3]
 RP SEQUENCE OF 502-1937 FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=90235862; PubMed=1691980;
 RA Boder E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
 RA Arnold H.H.;
 RT "Identification of three developmentally controlled isoforms of human
 myosin heavy chains.";
 RL Eur. J. Biochem. 189:55-65(1990).
 RN [4]
 RP SEQUENCE OF 860-1937 FROM N.A.
 RX MEDLINE=89234168; PubMed=2715179;
 RA Peghail R., Leinwand L.A.;
 RT "Molecular genetic characterization of a developmentally regulated
 human perinatal myosin heavy chain.";
 RL J. Cell Biol. 108:1791-1797(1989).
 RN [5]
 RP SEQUENCE OF 1-46 FROM N.A.
 RA Esser K., Tichar A., Myszkowski M.;
 RT "Isolation and characterization of the human perinatal MYC promoter.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (RLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M36769; AAC17185.1; -
 DR EMBL: Z38133; CAAB6293.1; -
 DR EMBL: X51592; CAAB3594.1; -
 DR EMBL: M35250; AAA36346.1; -
 DR EMBL: AF067143; AAC21557.1; -
 DR PIR: A30220; A30220.
 DR HSSP: P13538; 2MYS.
 DR MIM: 160741; -
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.

DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PSS0096; IQ; 1.
 KM Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Methylation; Alkylation; Multigene family;
 KM Calmodulin-binding.
 FT DOMAIN 1 780 MYOSIN HEAD-LIKE.
 FT DOMAIN 781 813 IQ.
 FT DOMAIN 842 1937 COILED COIL (POTENTIAL).
 FT NP_BIND 181 188 ATP.
 FT DOMAIN 658 680 ACTIN-BINDING.
 FT DOMAIN 760 774 ACTIN-BINDING.
 FT MOD_RES 132 132 METHYLATION (TR-) (POTENTIAL).
 FT MOD_RES 698 698 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 708 708 ALKYLATION (SH-2) (POTENTIAL).
 FT CONFLICT 15 15 A -> R (IN REF. 2).
 FT CONFLICT 970 970 E -> Q (IN REF. 1 AND 4).
 FT CONFLICT 1072 1072 M -> N (IN REF. 3).
 FT CONFLICT 1247 1247 N -> H (IN REF. 1 AND 4).
 FT CONFLICT 1251 1252 MC -> DGG (IN REF. 3).
 FT CONFLICT 1261 1261 E -> G (IN REF. 1 AND 4).
 FT CONFLICT 1287 1287 K -> Q (IN REF. 1 AND 4).
 FT CONFLICT 1377 1378 K -> Q (IN REF. 3).
 FT CONFLICT 1504 1505 EN -> AH (IN REF. 1 AND 4).
 FT CONFLICT 1847 1847 E -> D (IN REF. 1 AND 4).
 FT CONFLICT 1914 1914 D -> H (IN REF. 2).
 SQ SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;

Query Match 9.3%; Score 168; DB 1; Length 1937;
 Best Local Similarity 23.2%; Pred. No. 1.7;
 Matches 88; Conservative 70; Mismatches 166; Indels 56; Gaps 14;

QY 15 NSADESVKPNLTETSKKTTDSNAVLAVK-----EVALLSIDELAKIKKINDGS 69
 DB 914 NKIQLEAKIKETVERAEEREEIEMALTKARKLEDCSELEKIDIDLETLAKEREKHA 973
 QY 70 LNEARNRNSLAGATITSLITQKSLKNGSGLEKTIKAAKCCSEFSTKIKDNHQL 129
 DB 974 TEKKVKN-----LTEEMA-----GLOETIA--KLSE--KKAQETHQQT 1009
 QY 130 GIGCVTDENAKKAILKANAAGKDKGVELEKLSLESLSKAKEMLVANSKELTSPVH 189
 DB 1010 LDDLOAEKQVNLTKA-----KTKLEQYVDLGGSLGQEKRLMD--LERAKKRLGGLK 1064
 QY 190 GNNSGKDGNTSANSADSV--KGNPLTETSKITTESNAVVLAV--KEYETLLTSIDEL-- 243
 DB 1065 AGESTMDMENDKQDLEKLEKFEISNLISKIEDQAVETIQOKIKIKELQARIEELGEE 1124
 QY 244 -----AKAIGKIKNDVSLD--NEADHNGSLIGAVLISLTKIKISAIKDSGLAKAIE 296
 DB 1125 IEAERASRAKAEKORSDLSRELEIEISERLEEAGATSAQVELKRRBA--EFGKLRDLE 1182
 QY 297 KAKKCEEFPAKJKEGHTD--LGKEGYTDNAKKAAILKTNNDKTKGADLEKLEPESVKNLS 355
 DB 1183 EATLQHEAMVAALKKHADSMAELGECIDMLQRYKOKLEKSELMETDOLSNAPALS 1242
 QY 356 KAKEM-----LTNSVKEL 369
 DB 1243 KAKGNLEKKMCRSLDQVSEL 1262

RESULT 13
 MYSB.CAEEL
 ID MYSB.CAEEL STANDARD; PRT; 1966 AA.
 AC P02566;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 21-JUL-1986 (rel. 01, Last sequence update)
 DT 20-AUG-2001 (rel. 40, Last annotation update)
 DE MYOSIN HEAVY CHAIN B (MHC B).
 GN UNC-54 OR MYO-4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 myosin heavy chain gene are not separated by introns."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RN [2]
 RP SEQUENCE OF 850-1966 FROM N.A.
 RX MEDLINE=8272395; PubMed=720124;
 RA McLachlan A.D., Karn J.;
 RT "Periodic charge distributions in the myosin rod amino acid sequence
 match cross-bridge spacings in muscle."
 RL Nature 299:226-231(1982).
 RN [3]
 RP SEQUENCE OF 1876-1966 FROM N.A.
 RX MEDLINE=83232892; PubMed=6571695;
 RA Willis N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
 RA Waterston R.H.;
 RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber
 nonsense mutations via altered transfer RNA."
 RL Cell 33:575-583(1983).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
 ELEGANS.
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: J01050; AAA28124.1; -
 DR EMBL: V01494; CAA24738.1; -
 DR PIR: A02992; MKMK
 DR HSSP: P08799; 1MND
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head.1.
 DR Pfam: PF01576; Myosin_tail.1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head.1.
 DR SMART: SM00242; MYSC; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
 FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).
 FT DOMAIN 1165 1176 HINGE.
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMW).
 FT NP_BIND 177 184 ATP (BY SIMILARITY).
 FT DOMAIN 665 687 ACTIN-BINDING.
 FT DOMAIN 769 783 ACTIN-BINDING.

```

FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 705 705 ALKYLATION (SH-1).
FT MOD_RES 715 715 ALKYLATION (SH-2).
FT CONFLICT 1337 1337 E -> R (IN REF. 2).
FT CONFLICT 1880 1880 I -> L (IN REF. 2).
SQ SEQUENCE 1966 AA; 225125 MW; B66F0BB2FE27B67F CRC64;

Query Match 9.2%; Score 167; DB 1; Length 1966;
Best Local Similarity 22.8%; Pred.No.1.9; Mismatches 167; Indels 82; Gaps 19;
Matches 97; Conservative 79;

OY 8 KDGNTSANSADSVKGP-----NLTEISKITDSNAVLAVKEVALLS 51
DB 1121 KQDQSRISLEELLENERSKADAKSLDQRELELEDEKDEGCGAIAQVEVAKKE 1180
OY 52 SIDELAKAIGKTKKNDSDNEANN-ESLAGAVTIST---LITQKSLKNGSEGLE 106
DB 1181 A--ELA---KLIRD---LEENNMHNQGLGRKKHDAVAELTDQLDQINKAKAYE 1230
OY 107 KIAA-----AKCSEEFSTKLKDNHAGL-----ICQVTDENAKKAILKANA 148
DB 1231 KKAQAVRAEDLAQDLDETSGL--NNEKLAQELDLTQSKADQSKQLDQDFTSL 1288
OY 149 AGK-----DKGYEIEKLSGSLESISKAKEMLANSVKELTSPVHGNS 193
DB 1289 KRLHSENGDLVROLDEADSVQVQLRLKQSLQLEEARPTADEPARERQVAAQAKNY 1348
OY 194 GKDGNTSANSADSVKGPPLTETSKITETESNAVLAVK---EVEITLISIDELAKAIGK 250
DB 1349 QHEADQLOSLSELEIEGKN--ETLQKLSKANADIQQWKARFEGEGLKNA-DELEDAKRRQ 1405
OY 251 IK--NDV--SLDNEADHNGSL-ISCAYLISNLITKKAISAKDSGEIKAEIEKAKCSEEF 305
DB 1406 AAKINELQALDAANSKNSLSEKTSRLVGDLDAQVDERANGVASA-LEKKQGFPEKI 1464
OY 306 TAKLGEHDLDEKGVTDNNAKALKITNTKAGDELEKLFESVKNLISKAKEMLTNS 365
DB 1465 IDEMKRKTDDLAEE--LDGAQRDLRNTSTDLFKAKNAOEELAEVVEGLRRENKS-LSQE 1520
OY 366 VKELT 370
DB 1521 IKDLT 1525

RESULT 14
HUMAN STANDARD: PRT: 1713 AA.
LMA3 HUMAN 016787; 013679; 013680.
PRT 01-NOV-1997 (Rel. 35, Created)
PRT 01-NOV-1997 (Rel. 35, Last sequence update)
PRT 20-NOV-2001 (Rel. 40, Last annotation update)
DE LAMININ ALPNA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).
GN LAMA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Keratinocytes;
RX MEDLINE=94357926; PubMed=8077230;
RA Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT "Cloning of the lama3 gene encoding the alpha 3 chain of the adhesive
RT ligand epiligrin. Expression in wound repair.";
RL J. Biol. Chem. 269:22779-22787(1994).
RN [2]
RP SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=96163880; PubMed=8586427;
RA Vidal F., Baudouin C., Miguel C., Galliano M.-F., Christiano A.M.,
RA Vito J., Ortonne J.-P., Meneguzzi G.;
RT "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT of a homozygous deletion in a patient with Herlitz junctional

```

```

RT epiligrin/bullousa."
RL Genomics 30:273-280(1995).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC 6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC KERATINOCYTES.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC AND LAMININ-7 (KS-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -1- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
CC A.
CC -1- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC AND EPITHELIAL FOCAL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC IN THE CNS.
CC -1- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAIN G IS GLOBULAR.
CC -1- DISEASE: DEFECTS IN LAMA3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS
CC BULLOSA (JEB) GRAVIS, ALSO KNOWN AS HERLITZ TYPE JEB. IT IS A
CC BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION
CC OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED
CC NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT
CC MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS
CC BULLOSA.
CC -1- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L34155; AAA59483.1; -
DR EMBL: X85107; CA559428.1; -
DR EMBL: X85108; CA559429.1; -
DR HSSP: P02468; ITLE.
DR MIM: 600805; -
DR MIM: 226700; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR003129; TSPN.
DR Pfam: PF00053; laminin_EGF_2.
DR Pfam: PF00054; laminin_EGF_2.
DR SMART: SM00180; EGF_Lam; 2.
DR SMART: SM00282; LamG; 2.
DR SMART: SM00210; TSPN; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01248; LAMININ_TYPE_EGF_2.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW Alternative splicing; Epidermolysis bullosa.
FT SIGNAL 1
FT SIGNAL 20
FT POTENTIAL.

```

FT	CHAIN	21	1713	LAMININ ALPHA-3 CHAIN.
FT	DOMAIN	46	201	DOMAIN III A.
FT	DOMAIN	67	185	2-5 X LAMININ EGF-LIKE REPEATS.
FT	DOMAIN	67	113	LAMININ EGF-LIKE 1.
FT	DOMAIN	114	166	LAMININ EGF-LIKE 2.
FT	DOMAIN	167	185	LAMININ EGF-LIKE 3 (INCOMPLETE).
FT	DOMAIN	202	793	DOMAIN II AND I (HEPAT REPEATS).
FT	DOMAIN	794	1713	5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT	DOMAIN	794	970	LAMININ G-LIKE 1.
FT	DOMAIN	971	1139	LAMININ G-LIKE 2.
FT	DOMAIN	1140	1353	LAMININ G-LIKE 3.
FT	DOMAIN	1354	1529	LAMININ G-LIKE 4.
FT	DOMAIN	1530	1713	LAMININ G-LIKE 5.
FT	DOMAIN	231	327	COILED COIL (POTENTIAL).
FT	DOMAIN	396	548	COILED COIL (POTENTIAL).
FT	DOMAIN	594	621	COILED COIL (POTENTIAL).
FT	DOMAIN	702	765	COILED COIL (POTENTIAL).
FT	DOMAIN	1686	1713	COILED COIL (POTENTIAL).
FT	DISULFID	67	76	BY SIMILARITY.
FT	DISULFID	69	83	BY SIMILARITY.
FT	DISULFID	86	95	BY SIMILARITY.
FT	DISULFID	98	111	BY SIMILARITY.
FT	DISULFID	114	126	BY SIMILARITY.
FT	DISULFID	116	135	BY SIMILARITY.
FT	DISULFID	137	146	BY SIMILARITY.
FT	DISULFID	149	164	BY SIMILARITY.
FT	DISULFID	202	202	INTERCHAIN (PROBABLE).
FT	DISULFID	205	205	INTERCHAIN (PROBABLE).
FT	SITE	658	660	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBOHYD	542	542	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	645	645	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	745	745	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	882	882	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	964	964	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1108	1108	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1131	1131	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1325	1325	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1477	1477	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1667	1667	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	VARSPLIC	1	45	MMWLTFFGALDQCCLGYSQQOQVRPFLOPQSOQASVBE
FT				FRPS -> KVSSYGGLTYKQSFALPDWNLKRRKDDVO
FT				TCQHSMSIYEETNTPRPDLRHLHGHRVHVEGFRASSRAV
FT				SREELMTVYSRLADVRIGQLGTFETFORLTSEVLEFASD
FT				GGRIALAVETICACDPAYAGSC (IN ISOFORM B).
FT				W -> R (IN REF. 2).
FT				AVG -> GMC (IN REF. 2).
FT				M -> K (IN REF. 2).
FT				R -> L (IN REF. 2).
FT				E -> O (IN REF. 2).
FT				D -> A (IN REF. 2).
FT				G -> A (IN REF. 2).
FT				1713 AA: 189304 MW: 45EABBE101B60D3 CRC64:

Query Match	9.2%	Score 166.5	DB 1	Length 1713
Best Local Similarity	23.0%	Pred. No. 1.7		
Matches	90	Conservative	71	Mismatches 138; Indels 93; Gaps 16
Qy	9	DGNTSANSADSVKGGPNLTETSKITSDSNAVU----	LAVEKEVALLSSIDELAKAIGKRT	64
Db	396	ENNGLANSTIRDS-----LNEYEAKLSDLRARLOEAAQAQKANGNLQENBERALGATQGV		450
Qy	65	KNDGSLDNEANNESILAGATTISLTITQKLSKLNQSEGLKRIKIAA--KKCSFEFTKJK		123
Db	451	KEINSLQSFYTYLTAAOSSLTQETNIALQIMKK--SOKYEKLAASLAEAOEISDKYR		507
Qy	124	DHHAOLGIGYDENAKKAAILKANAGKRGVDELEKISGSLESLAKEMLAN--SVKE		182
Db	508	E-----LSRAGKGSUVEAEKAKHARSQLOELAKOLOEIKIRNASSGDE		547
Qy	183	LN-----SPVYHGNGSGKDGNTSANSADSVKGGPNLTETSKITTESNAVUAVLAVE		232
Db	548	LVRCAVDATATAYENTLNLAKAEDNARMASSSES-----ALQTVIKE		590

Oy	223	----	VEILLINGSIDEL--KAIGKIRINDVS--LDN-EADHONSLSJGATVLSLITRK	281
Dd	501	DLPRAKATLLSSNSDKLLEAKKTOKKLQOEVSPLANNIQQTLNLIVTYOKEVIDRNLTLR	650	
Oy	282	-----ISAIKDSGELAIEERAKKCSEEFYAKLKGERTDGRKGVD-----D	324	
Dd	651	DGLHGICRGDIAMDISSA--KSMVKRKANDIDEVLGDINPIQTQDV--BRIDFTVGRTONE	706	
Oy	325	NAKKAILETKNDNKPTGADELEKLFESVYNLSK	356	
Dd	707	DFKKALTDADNSVNKLTLPDLMKRIESSINO	738	
RESULT 15				
ID	RBP1_PLAVB	STANDARD:	PRT; 2869 AA.	
AC	000798:			
Dt	01-APR-1993	(Rel. 25, Created)		
Dt	01-APR-1993	(Rel. 25, Last sequence update)		
Dt	01-OCT-1996	(Rel. 34, Last annotation update)		
De	RETICULOCYTE BINDING PROTEIN 1 PRECURSOR.			
Gn	RBP1.			
Os	Plasmodium vivax (strain Belen).			
Oc	Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.			
Rn	NCHI_TaxID=31273;			
Rx	[1]			
Rp	SEQUENCE FROM N.A.			
Rx	MEDLINE=92315338; PubMed=1617731;			
Ra	Gallinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;			
Rt	"A reticulocyte-binding protein complex of Plasmodium vivax merozoites."			
Rl	Cell 61:1213-1226(1992).			
Cc	-1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO			
Cc	HUMAN RETICULOCYTE CELLS.			
Cc	-1- SUBUNIT: HOMODIMER (POTENTIAL).			
Cc	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.			
Cc	-----			
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration			
Cc	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
Cc	the European Bioinformatics Institute. There are no restrictions on its			
Cc	use by non-profit institutions as long as its content is in no way			
Cc	modified and this statement is not removed. Usage by and for commercial			
Cc	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
Cc	or send an email to license@isb-sib.ch).			
Cc	-----			
Dc	EMBL; M88097; AAA29743.1; .			
Dr	HSPD; P36956; IAM9.			
Kw	Malaria; Receptor; Signal; Transmembrane.			
Ft	SIGNAL	1	17	
Ft	CHAIN	18	2869	
Ft	DOMAIN	18	2807	
Ft	TRANSMEM	2808	2826	
Ft	DOMAIN	2827	2869	
Ft	SITE	1030	1032	
Ft	SITE	2399	2601	
Se	SEQUENCE	2869 AA; 330213 MW; B9DBEA42205EBCF CRC64;		

```

Query Match Similarity      9.1%; Score 165.5; DB 1; Length 2869;
Best Local Similarity      22.4%; Pred No. 3.4;
Matches 101; Conservative 63; Mismatches 124; Indels 163; Gaps 20.

OY 25 NLTETSKITT-DSNAVLAARVEEALLSSIDELANAIKRIKNDGSLDNEANRRESLLAG 83
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1141 HINTVAAGHTSKRNKELLISVKEVEERKLNLEVO--NEDYKKRVNP--ENE----- 1185
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 84 AVTISTLTQKTSKNGS-EGLEKRIIAAKKSEESTSTLKNHQAOLGQVTDENAKKA 142
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1186 -----KOLEAIRGNSKSLLEYI-----NKH-----VSEMTQLESTYN 1217
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 143 IUKANAAGKDR--GYEDELKLSGL-----ESLSKAAREMLANSVVELTSPVHGNSGCD 196
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```
Db 1218 TLKSNAGKRENEHDEELNKTGQMRDIYEKIKKIAEELKGTYNEL-----KD 1266
Oy 197 GNTSANSADSVKGPNTLEISKIT-ESNAVYLAKEVEETLLTSIDELAK-----245
    | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1267 ANEKANKVPEPERNIGHVLERITVEKDKAGKVEEMNSLKTIEKLIQETSDSQNEL 1326
Oy 246 ---AIGKKIKN-----DVSLDNEAD-----HNGSLISGA 271
    | | : : : : : : : : : : : : : : : : : : : : : : : :
Db 1327 VTTSTIKHLENAGYEDVTKRNEEDSIQIREKAKSLETLDENKKLVQOVNMNLQSAIQN 1386
Oy 272 YLISN-----LITKKISAI-----KDSGELK---AE 294
    | | : : : : : : : : : : : : : : : : : : : : : : : :
Db 1387 AGISKELNELKGVIELLISTNWSIIEYKKNSSSVRESOLANGELFKAEGEKKNASAR 1446
Oy 295 IEKAKKSESEFTAKLKEHTDLGKEGYTDDNAKKAALK-----TNNDKTK-----339
    | | : : : : : : : : : : : : : : : : : : : : : : : :
Db 1447 LAEAEKLEQIVKDL--DYSDIDDVKKIEGIRKREILKMKESALTFWEESEKFKQMCSSH 1504
    | | : : : : : : : : : : : : : : : : : : : : : : : :
    340 --GADELEKLFESVKKLSKAKAKEMLTNSVKE 368
    | | : : : : : : : : : : : : : : : : : : : : : : : :
Db 1505 MENAKEGKKRIEYLNKNGDGKANITDSOME 1535
```

Search completed: March 18, 2002, 10:11:52
Job time: 970 sec

QY 181 KELTSPPV 188
 DB 187 KELTSPPV 194

RESULT 2

044977 PRELIMINARY; PRT: 211 AA.
 AC 044977;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C.
 GN OSCPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PARE:
 RX MEDLINE=95395018; PubMed=7655660;
 RA Jauris-Helpe S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,
 RA Soutschek E., Will G., Wilske B.;
 RT "Molecular analysis of genes encoding outer surface protein C (OspC)
 RT of Borrelia burgdorferi sensu lato: relationship to ospA genotype and
 RT evidence of lateral gene exchange of ospC.";
 RL J. Clin. Microbiol. 33:1860-1866(1995).
 DR EMBL; X81522; CA57242.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 211 AA; 22214 MW; BFE8C7820CDE10D3 CRC64;

Query Match 49.0%; Score 886; DB 2; Length 211;
 Best Local Similarity 97.3%; Pred. No. 1.7e-32;
 Matches 183; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNNGSKDNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAI 60
 DB 17 ISCNNGSKDNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAI 76
 QY 61 GKIKNDGSLDNEANRNESILAGAVTISTLITOKLSKNGSEGLKETAARKCEEST 120
 DB 77 GKIKNDGSLDNEANRNESILAGAVTISTLITOKLSKNGSEGLKETAARKCEEST 136
 QY 121 KIKDNHQAQIGQVTDENAKKAILKANAAGKDGVEELEKLSGLESLSKAKEMLANSV 180
 DB 137 KIKDNHQAQIGQVTDENAKKAILKANAAGKDGVEELEKLSGLESLSKAKEMLANSV 196
 QY 181 KELTSPPV 188
 DB 197 KELTSPPV 204

RESULT 3
 031120 PRELIMINARY; PRT: 191 AA.
 AC 031120;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OC10;
 RA Wang L.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL SubMITTED (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF029869; AAB86552.1;

DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 191 AA; 20502 MW; E4FB56BCH61740F9 CRC64;

Query Match 48.9%; Score 885; DB 2; Length 191;
 Best Local Similarity 100.0%; Pred. No. 1.7e-32;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 NNSGKDNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEALLSIDELAKAIGK 250
 DB 10 NNSGKDNTSANSADSVKGPMLTEISKRTTESNAVLAVKEVEALLSIDELAKAIGK 69
 QY 251 IKNDVSLDNEADHNSLISGAVTISNLTTRKISAIGSGELKAEIERAKKCSEESTAKL 310
 DB 70 IKNDVSLDNEADHNSLISGAVTISNLTTRKISAIGSGELKAEIERAKKCSEESTAKL 129
 QY 311 GEHTDLGREGVTDNNAKKAIIKTNNDKTKGADLEKLFESYKNSKAKEMLTNSVKELT 370
 DB 130 GEHTDLGREGVTDNNAKKAIIKTNNDKTKGADLEKLFESYKNSKAKEMLTNSVKELT 189
 QY 371 S 371
 DB 190 S 190

RESULT 4
 P94242 PRELIMINARY; PRT: 193 AA.
 AC P94242;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZS7;
 RX MEDLINE=96296448; PubMed=8709845;
 RA Livey I., Gibbs C.P., Schuster R., Dörner F.;
 RT "Evidence for lateral transfer and recombination in OspC variation in
 RT Lyme disease Borrelia.";
 RL Mol. Microbiol. 18:257-269(1995).
 DR EMBL; I42868; AAB37011.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 193 AA; 20189 MW; 64E30C240E5F54CA CRC64;

Query Match 48.9%; Score 884; DB 2; Length 193;
 Best Local Similarity 98.4%; Pred. No. 1.9e-32;
 Matches 183; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CNNSGKDNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAIGK 62
 DB 1 CNNSGKDNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEVEALLSIDELAKAIGK 60
 QY 63 KIKNDGSLDNEANRNESILAGAVTISTLITOKLSKNGSEGLKETAARKCEESTK 122
 DB 61 KIKNDGSLDNEANRNESILAGAVTISTLITOKLSKNGSEGLKETAARKCEESTK 120
 QY 123 KDNHQAQIGQVTDENAKKAILKANAAGKDGVEELEKLSGLESLSKAKEMLANSV 182
 DB 121 KDNHQAQIGQVTDENAKKAILKANAAGKDGVEELEKLSGLESLSKAKEMLANSV 180

OY 183 LTSPV 188
 DB 181 LTSPV 186

RESULT 5

Q9S3P2 5 PRELIMINARY: PRT: 192 AA.
 ID Q9S3P2
 AC Q9S3P2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID-139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-297;
 RA MEDLINE-96296448; PubMed-8709845;
 RA Livey I., Gibbs C.P., Schuster R., Dorner F.;
 RT "Evidence for lateral transfer and recombination in OSpC variation in
 RT Lyme disease Borrelia.";
 RL Mol. Microbiol. 18:257-269(1995).
 DR EMBL; L42893; AAB37001.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 192
 SQ SEQUENCE 192 AA; 20472 MW; 46AC8F93E4DEED6C CRC64;

Query Match 48.6%; Score 879; DB 2; Length 192;
 Best Local Similarity 99.4%; Pred. No. 3.1e-32;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 191 NNSGKDGNTSANSADSVKGNLTKTESNAVLAVKEVETLLTSIDELAKAIGKK 250
 DB 2 NNSGKDGNTSANSADSVKGNLTKTESNAVLAVKEVETLLTSIDELAKAIGKK 61
 OY 251 INNDVSLDNEADHNGSLISGAVLISNLTIRKISAIKDSGLKAEIKAKKCEEFPAKLK 310
 DB 62 INNDVSLDNEADHNGSLISGAVLISNLTIRKISAIKDSGLKAEIKAKKCEEFPAKLK 121
 OY 311 GEHTDLGKEGVTDNNKKAIIKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 370
 DB 122 GEHTDLGKEGVTDNNKKAIIKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 181
 OY 371 S 371
 DB 182 S 182

RESULT 6

Q44719 6 PRELIMINARY: PRT: 210 AA.
 ID Q44719
 AC Q44719;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OSpC.
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID-139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HB19;
 RA MEDLINE-95154673; PubMed-7651744;
 RA Stevenson B., Barthold S.W.;

RT "Expression and sequence of outer surface protein C among North
 RT American isolates of Borrelia burgdorferi.";
 RL FEMS Microbiol. Lett. 124:367-372(1994).
 DR EMBL; U04281; AAC43297.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 210 AA; 22526 MW; 2D672991D584E4EE CRC64;

Query Match 48.6%; Score 879; DB 2; Length 210;
 Best Local Similarity 99.4%; Pred. No. 3.4e-32;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 191 NNSGKDGNTSANSADSVKGNLTKTESNAVLAVKEVETLLTSIDELAKAIGKK 250
 DB 20 NNSGKDGNTSANSADSVKGNLTKTESNAVLAVKEVETLLTSIDELAKAIGKK 79
 OY 251 INNDVSLDNEADHNGSLISGAVLISNLTIRKISAIKDSGLKAEIKAKKCEEFPAKLK 310
 DB 80 INNDVSLDNEADHNGSLISGAVLISNLTIRKISAIKDSGLKAEIKAKKCEEFPAKLK 139
 OY 311 GEHTDLGKEGVTDNNKKAIIKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 370
 DB 140 GEHTDLGKEGVTDNNKKAIIKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 199
 OY 371 S 371
 DB 200 S 200

RESULT 7

Q08138 7 PRELIMINARY: PRT: 203 AA.
 ID Q08138
 AC Q08138;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
 GN OSpC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID-139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DK7;
 RA MEDLINE-94075528; PubMed-8253951;
 RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
 RT "Polymorphism in OSpC gene of Borrelia burgdorferi and
 RT immunoreactivity of OSpC protein: implications for taxonomy and for
 RT use of OSpC protein as a diagnostic antigen";
 RL J. Clin. Microbiol. 31:2570-2576(1993).
 CC -1 SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.

DR EMBL; X73625; CAA52004.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 KW Signal; Lipoprotein; Outer membrane.
 FT SIGNAL 1 18
 FT CHAIN 19 203
 FT NON_TER 203
 SQ SEQUENCE 203 AA; 21364 MW; AC9B29FE0E69B48F CRC64;

Query Match 47.7%; Score 862; DB 2; Length 203;
 Best Local Similarity 95.7%; Pred. No. 1.9e-31;
 Matches 179; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MACNNSGKDGNTSANSADSVKGNLTKTESNAVLAVKEVETLLTSIDELAKAI 60
 DB 17 ISCNNSGKDGNTSANSADSVKGNLTKTESNAVLAVKEVETLLTSIDELAKAI 76

OY 61 GKIKNDGSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAARKCSEEFST 120
 DB 77 GKIKNDGSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAARKCSEEFST 136
 OY 121 KIKDNHQAQIGQVTDENAKKAILKANAAAGDKGVEELEKLSGLESLSKRAKEMLANSV 180
 DB 137 KIKDNHQAQIGQVTDENAKKAILKANAAAGDKGVEELEKLSGLESLSKRAKEMLANSV 196
 OY 181 KETSPV 187
 DB 197 KSLQSYV 203

RESULT 8
 O44980 PRELIMINARY: PRT: 178 AA.
 ID 044980:
 AC 044980:
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BUR:
 RX MEDLINE=95286481; PubMed=7768799;
 RA Thøgersen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 Hansen K.;
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC";
 RL J. Bacteriol. 177:3036-3044(1995).
 DR EMBL; X84765; CA559236.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT 178
 SO SEQUENCE 178 AA; 18724 MW; 49A6F785145CC208 CRC64;

Query Match 46.9%; Score 849; DB 2; Length 178;
 Best Local Similarity 99.4%; Pred. No. 6e-31;
 Matches 177; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 11 NTSANSADESYKGNLPTFISKRTDSNAVLLAVEKVEALLSSIDELAKAIGKTIKNGSL 70
 DB 1 NTSANSADESYKGNLPTFISKRTDSNAVLLAVEKVEALLSSIDELAKAIGKTIKNGSL 60
 OY 71 DNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAARKCSEEFSTKIKDNHQAQ 130
 DB 61 DNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAARKCSEEFSTKIKDNHQAQ 120
 OY 131 IQGVTDENAKKAILKANAAAGDKGVEELEKLSGLESLSKRAKEMLANSVKELTSPV 188
 DB 121 IQGVTDENAKKAILKANAAAGDKGVEELEKLSGLESLSKRAKEMLANSVKELTSPV 178

RESULT 9
 O9R7B3 PRELIMINARY: PRT: 174 AA.
 ID 09R7B3:
 AC 09R7B3:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT)
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;

RX MEDLINE=97478003; PubMed=9336916;
 RA Ras N.M., Postic D., Foretz M., Baranton G.;
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
 RT U.S.A.'?";
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;
 RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U91795; AAB81892.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT 174
 SO SEQUENCE 174 AA; 18291 MW; A14F5B8BA151A278 CRC64;

Query Match 43.7%; Score 790; DB 2; Length 174;
 Best Local Similarity 97.0%; Pred. No. 2.3e-28;
 Matches 162; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MACNNSGKDGNTSANSADSYKGNLPTFISKRTDSNAVLLAVEKVEALLSSIDELAKAI 60
 DB 8 ISCNNSGKDGNTSANSADSYKGNLPTFISKRTDSNAVLLAVEKVEALLSSIDELAKAI 67
 OY 61 GKIKNDGSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAARKCSEEFST 120
 DB 68 GKIKNDGSLDNEANRNESLAGAVTTSTLTOKLSKNGSEGLKEKIAAARKCSEEFST 127
 OY 121 KIKDNHQAQIGQVTDENAKKAILKANAAAGDKGVEELEKLSGLES 167
 DB 128 KIKDNHQAQIGQVTDENAKKAILKANAAAGDKGVEELEKLSGLES 174

RESULT 10
 O9AGB1 PRELIMINARY: PRT: 165 AA.
 ID 09AGB1:
 AC 09AGB1:
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 GN OSPC-MODIFIED.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HB19;
 RA Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
 RA Lawson C.L., Swaminathan S.;
 RT "Crystal structure of outer surface protein C (OspC) from the Lyme
 RT disease spirochete, Borrelia burgdorferi";
 RL EMBO J. 0:0-0(2001).
 DR EMBL; AF337548; AAK21289.1; -;
 FT VARIANT 61
 FT 165
 SO SEQUENCE 165 AA; 17833 MW; 56E01536D22F61BF CRC64;

Query Match 43.4%; Score 786; DB 2; Length 165;
 Best Local Similarity 98.2%; Pred. No. 3.2e-28;
 Matches 161; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 208 VKPNTFISKRTTESNAVLLAVEKVEETLTSTIDELAKAIGKIKNDVSLDNEADHNGSL 267
 DB 1 MKPNTFISKRTTESNAVLLAVEKVEETLTSTIDELAKAIGKIKNDVSLDNEADHNGSL 60
 OY 268 ISGAVILSNLTIKKISAKDSGELKAEIEKAKCSEEFSTAKLGEHTDGLKGSVTDNNAK 327
 DB 61 MSGAVILSNLTIKKISAKDSGELKAEIEKAKCSEEFSTAKLGEHTDGLKGSVTDNNAK 120

OY 328 KALKTNNDKTKGADELEKLFESVKNLSKAKEMLTNSVKELTS 371
 DB 121 KALKTNNDKTKGADELEKLFESVKNLSKAKEMLTNSVKELTS 164

RESULT 11
 ID 034121 PRELIMINARY: PRT: 168 AA.

AC 034121: PRELIMINARY: PRT: 168 AA.
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;
 RX MEDLINE=97478003: PubMed=9336916;
 RA Ras N.M., Postic D., Foretz M., Baranton G.;
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
 U.S.A.'?";
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TETS;
 RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U91794; AAB81891.1; -;
 DR InterPro: IPR001800: Lipoprotein_6.
 DR Pfam: PF01441: Lipoprotein_6; 1.
 DR ProDom: PD001149: Lipoprotein_6; 1.
 FT NON_TER 168 168
 FT SEQUENCE 168 AA; 17704 MW; CDA45DD8A9C4409 CRC64;

Query Match 43.2%; Score 782; DB 2; Length 168;
 Best Local Similarity 98.2%; Pred. No. 5e-28;
 Matches 164; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 19 ESKGNGNLTEISKRTDSNAVLLAVEVEALLSIDELAKIKRTKNDSDNEANRNE 78
 DB 1 ESKGNGNLTEISKRTDSNAVLLAVEVEALLSIDELAKIKRTKNDSDNEANRNE 60
 OY 79 SLTAGVTTSTLTITOKLSKNGSEGLKEKTAIAKKCEEFSTKLDNHAQLGIGVTDEN 138
 DB 61 SLTAGVTTSTLTITOKLSKNGSEGLKEKTAIAKKCEEFSTKLDNHAQLGIGVTDEN 120
 OY 139 AKKAIKANAAAGKDGVEELEKLSGLESLSKAKEMLANSVKELTS 185
 DB 121 AKKAIKANAAAGKDGVEELEKLSGLESLSKAKEMLANSVKELTS 167

RESULT 12
 ID 031115 PRELIMINARY: PRT: 193 AA.
 AC 031115:
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-OC3;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF029862; AAB86545.1; -;

DR InterPro: IPR001800: Lipoprotein_6.
 DR Pfam: PF01441: Lipoprotein_6; 1.
 DR ProDom: PD001149: Lipoprotein_6; 1.
 FT NON_TER 193 193
 FT SEQUENCE 193 AA; 20596 MW; 018A4CB310475A58 CRC64;

Query Match 39.8%; Score 720.5; DB 2; Length 193;
 Best Local Similarity 79.8%; Pred. No. 9e-25;
 Matches 150; Conservative 16; Mismatches 21; Indels 1; Gaps 1;

OY 1 MACNNSGKDGNTSANSDESVEKGNLFEISKRTDSNAVLLAVEVEALLSIDELAKAI 60
 DB 7 ISCNNSGKDGNTSANSDESVEKGNLFEISKRTDSNAVLLAVEVEALLSIDELAKAI 66
 OY 61 GKKIKNDGSLDNEANRNEISLAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCEEFST 120
 DB 67 GKKIKNDVSLDNEADNNGSLIAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCEEFSTK 126
 OY 121 KLDNHAQLGIGVTDENAKKAIKANAAAGKDGVEELEKLSGLESLSKAKEMLANSV 180
 DB 127 KLEKHTDGLCKDQDTVAHAEALIKTNGT-KDKGALEKLFESVENLAKAEMLSNV 185
 OY 181 KELTSPVY 188
 DB 186 KELTSPVY 193

RESULT 13
 ID 031122 PRELIMINARY: PRT: 194 AA.
 AC 031122:
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-OC12;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF029871; AAB86554.1; -;
 DR InterPro: IPR001800: Lipoprotein_6.
 DR Pfam: PF01441: Lipoprotein_6; 1.
 DR ProDom: PD001149: Lipoprotein_6; 1.
 FT NON_TER 194 194
 FT SEQUENCE 194 AA; 20640 MW; 00A5E6E2D2CE0F7F CRC64;

Query Match 36.7%; Score 664; DB 2; Length 194;
 Best Local Similarity 75.1%; Pred. No. 9.1e-23;
 Matches 142; Conservative 19; Mismatches 26; Indels 2; Gaps 2;

OY 1 MACNNSGKDGNTSANSDESVEKGNLFEISKRTDSNAVLLAVEVEALLSIDELAKAI 59
 DB 7 ISCNNSGKDGNTSANSDESVEKGNLFEISKRTDSNAVLLAVEVEALLSIDELAKAI 66
 OY 60 IGKIKNDGSLDNEANRNEISLAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCEEFST 119
 DB 67 IGKIKNDGSLDNEANRNEISLAGAVTTSTLTITOKLSKNGSEGLKEKTAIAKKCEEFST 126
 OY 120 TKLDNHAQLGIGVTDENAKKAIKANAAAGKDGVEELEKLSGLESLSKAKEMLANSV 179
 DB 127 KLEGEHAQLGIENVTDENAKKAIKITDAA-KDKGALEKLFESVENLAKAEMLSNV 185
 OY 180 VKELTSPVY 188
 DB 186 VKELTSPVY 193

DB 186 VKELTSPV 194

RESULT 14

ID 034120 PRELIMINARY; PRT; 184 AA.

AC 034120;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-TETS;

RX MEDLINE=97478003; PubMed=9336916;

RA Ras N.M., Postic D., Foretz M., Baranton G.;

RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the U.S.A.'?";

RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN-TETS;

RA Marti-Ras N., Postic D., Foretz M., Baranton G.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U91793; AAB81890.1; "

DR InterPro: IPR001800; Lipoprotein_6.

DR Pfam: PF01441; Lipoprotein_6; 1.

DR Prodom: PD001149; Lipoprotein_6; 1.

FT NON-TER 1

FT NON-TER 184

SQ SEQUENCE 184 AA; 19391 MM; 981C12B79273FB0E CRC64;

Query Match 36.6%; Score 662; DB 2; Length 184;
Best Local Similarity 76.3%; Pred. No. 1e-22;
Matches 142; Conservative 15; Mismatches 27; Indels 2; Gaps 2;

QY 3 C NNSGKDGNTSANSADSVKGNLTETSKITDSNAVLLAVEVALLSIDELAKAIG 62
DB 1 C NNSGKDGNTSANSADSVKGNLTETSKITDSNAVLLAVEVALLSIDELAKAIG 60

QY 63 KIKNDGSLDNEANPNESLLAGAVTISLTITQKLSKNGSEGLKEKIAAKKSEEFSTKL 122
DB 61 KVEANG-LGNEDRRTSLLAGHETISLTITQKLTALKDSGGLKAEIARAKKCEAFYTKL 119

QY 123 KDNHAQIGIGVTDENAKKAILKANAAAGKDGVELEKLSGSLSELSRAKEMLANSVK 182
DB 120 KDNHAQIGIGVTDENAKKAILKLTNAT-KDKGAEDLEKLFRAVESLSKAQDALTNVYKE 178

QY 183 LTSPV 188
DB 179 LTNPV 184

RESULT 15

P94234 PRELIMINARY; PRT; 193 AA.

AC P94234;

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE OUTER SURFACE PROTEIN C (FRAGMENT).

OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OC Borrelia burgdorferi (Lyme disease spirochete).

OX NCBI_TaxID=139;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-28354;

RX MEDLINE=96296448; PubMed=8709845;

RA Livey I., Gibbs C.P., Schuster R., Dörner F.;

RT "Evidence for lateral transfer and recombination in OspC variation in Lyme disease Borrelia.";

RT Mol. Microbiol. 18:257-269(1995).

DR EMBL: L42895; AAB37003.1; "

DR InterPro: IPR001800; Lipoprotein_6.

DR Pfam: PF01441; Lipoprotein_6; 1.

DR Prodom: PD001149; Lipoprotein_6; 1.

FT NON-TER 1

FT NON-TER 193

SQ SEQUENCE 193 AA; 20411 MM; 05B68720F061E2A0 CRC64;

Query Match 36.6%; Score 662; DB 2; Length 193;
Best Local Similarity 75.9%; Pred. No. 1.1e-22;
Matches 142; Conservative 17; Mismatches 26; Indels 2; Gaps 2;

QY 3 C NNSGKDGNTSANSADSVKGNLTETSKITDSNAVLLAVEVALLSIDELAKAIG 61
DB 1 C NNSGKDGNTSANSADSVKGNLTETSKITDSNAVLLAVEVALLSIDELAKAIG 60

QY 62 KIKNDGSLDNEANPNESLLAGAVTISLTITQKLSKNGSEGLKEKIAAKKSEEFSTKL 121
DB 61 KIQNGGLAVEAGHNGTLLAGAVTISLTITQKLDGLKNSKELKEINAKKSEDFYTK 120.

QY 122 KDNHAQIGIGVTDENAKKAILKANAAAGKDGVELEKLSGSLSELSRAKEMLANSVK 181
DB 121 LEGBHAQIGIENVTDENAKKAILITDAA-KDKGAEDLEKLFRAVENLAKAKEMLANSVK 179

QY 182 LTSPV 188
DB 180 ELTSPV 186

Search completed: March 18, 2002, 10:10:54
Job time: 977 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:32 : Search time 118.14 Seconds
(without alignments)
230.734 Million cell updates/sec

Title: us-09-596-746a-38

Perfect score: 1797
Sequence: 1 MACNNSGKDGNGASNSADES.....ENLSKAQDTLKNVKELTLS 368

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : A.Geneseq.1101.*

```
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1797	100.0	368	22	AA62717
2	1789	99.6	368	22	AA62736
3	1533	85.3	367	22	AA62715
4	1530	85.1	391	22	AA62735
5	1485.5	82.7	368	22	AA62718
6	1482.5	82.5	392	22	AA62734
7	1479	82.3	369	22	AA62714
8	1476	82.1	393	22	AA62732
9	1225.5	68.2	377	22	AA62713
10	1224	68.1	373	22	AA62711
11	1221	67.9	397	22	AA62729

12	1220	67.9	369	22	AA62716	Borrelia sp chimera
13	1219.5	67.9	400	22	AA62739	Borrelia sp chimera
14	1217	67.7	393	22	AA62731	Borrelia sp chimera
15	1178	65.6	378	22	AA62712	Borrelia sp chimera
16	1176.5	65.5	374	22	AA62710	Borrelia sp chimera
17	1173.5	65.3	398	22	AA62728	Borrelia sp chimera
18	1172.5	65.2	370	22	AA62715	Borrelia sp chimera
19	1172	65.2	401	22	AA62738	Borrelia sp chimera
20	1169.5	65.1	394	22	AA62730	Borrelia sp chimera
21	1167.5	65.0	560	22	AA62724	Borrelia sp chimera
22	1157.5	64.4	384	22	AA62726	Borrelia sp chimera
23	1154.5	64.2	408	22	AA62737	Borrelia sp chimera
24	1142	63.6	378	22	AA62725	Borrelia sp chimera
25	1134	63.1	401	22	AA62733	Borrelia sp chimera
26	1099.5	61.2	410	22	AA62740	Borrelia sp chimera
27	1098.5	61.1	386	22	AA62727	Borrelia sp chimera
28	901	50.1	191	22	AA62705	Burgdorferi ospc
29	899	50.0	184	22	AA62702	Burgdorferi ospc
30	889	50.0	191	22	AA60884	Borrelia 28691 ospc
31	889	49.5	191	22	AA62706	Burgdorferi ospc
32	673	37.5	466	16	AA62740	B31 outer surface
33	663	36.9	193	22	AA62709	Burgdorferi ospc
34	657.5	36.6	212	16	AA62734	Outer surface prot
35	651.5	36.3	194	15	AA60885	Borrelia 2591 ospc
36	640	35.6	587	16	AA62746	B31 Osp-A/antigen
37	638.5	35.5	466	16	AA62739	B31 outer surface
38	623.5	34.7	192	22	AA62703	Burgdorferi ospc
39	623.5	34.7	210	16	AA61935	Outer surface prot
40	622.5	34.6	192	22	AA62721	Burgdorferi ospc
41	621.5	34.6	192	15	AA60886	Borrelia ip2 ospc
42	620.5	34.5	209	22	AA62720	Burgdorferi ospc
43	617.5	33.9	210	16	AA62727	B. burgdorferi str
44	608.5	34.4	178	15	AA62771	Borrelia 2591 anti
45	605.5	33.7	194	15	AA60894	Borrelia H9 antige

ALIGNMENTS

```
RESULT 1
ID AAB62717 standard; Protein: 368 AA.
XX
AC AAB62717;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein spq ID NO: 38.
XX
KM Borrelia: ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN W0200078966-A1.
XX
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000MO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX
XX (UANY ) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Dattwyler RJ, Sehnost G, Dykhuizen D, Luft RJ, Gomes-Solecki M;
XX WPI; 2001-050113/06.
XX N-PSDB; AAF29021.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
```

XX Claim 43; Page 97; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

XX Sequence 368 AA:

Query Match 100.0%; Score 1797; DB 22; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.5e-107;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MACNNSGKGNASANSADSVKGPNTLTKITSSNAVYLAKEVETLLASIDELATKA 60
DB 1 macnsgkgnasansadesvkgpnlteiskkltesnavylavkevetllasidelatka 60
OY 61 IGGKIGNNGLEANSOKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTEFTMK 120
DB 61 igkigngngleangskntsllsgayaissdlaeklnvlkneelkekidtakgscstefmk 120
OY 121 LKSEHAVLGLDNLTDNAORAILKKHANKDKGALEKLFKAVENTLSKAADTLKNAVKE 180
DB 121 lksehavlgldnltdnagrailkkhankdkgaalekrlfkavenlskaadtlknave 180
OY 181 LTPSVIHGNNRKGDNASTNSADESVKGPNTLTKITSSNAVYLAKEVETLLASIDE 240
DB 181 ltpsvihgnnsrkdgnastnsadesvkgpnlteiskkltesnavylavkevetllaside 240
OY 241 LATKAIKKIGKNGNGLEANSOKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTE 300
DB 241 latkaiigkigngngleangskntsllsgayaissdlaeklnvlkneelkekidtakgscst 300
OY 301 EFTNKLKSEHAVLGLDNLTDNAORAILKKHANKDKGALEKLFKAVENTLSKAADTLK 360
DB 301 eftnklksehavlgldnltdnagrailkkhankdkgaalekrlfkavenlskaadtlkn 360
OY 361 NAVKELTS 368
DB 361 navkeltss 368

RESULT 2

AAB62736
ID AAB62736 standard; Protein: 391 AA.

XX AAB62736;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 76.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN W0200078966-A1.

PD 28-DEC-2000.

PE 19-JUN-2000; 2000WO-US16915.

PA (UYNV) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29040.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause

PT Lyme disease are used to immunize animals and detect immune responses

PS Claim 43; Page 144-145; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

XX Sequence 391 AA:

Query Match 99.6%; Score 1789; DB 22; Length 391;
Best Local Similarity 99.7%; Pred. No. 5.1e-107;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACNNSGKGNASANSADSVKGPNTLTKITSSNAVYLAKEVETLLASIDELATKA 61
DB 2 acnnsghkgnasansadesvkgpnlteiskkltesnavylavkevetllasidelatka 61
OY 62 GKKIGNNGLEANSOKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTEFTMK 121
DB 62 gkkgngngleangskntsllsgayaissdlaeklnvlkneelkekidtakgscstefmk 121
OY 122 KSEHAVLGLDNLTDNAORAILKKHANKDKGALEKLFKAVENTLSKAADTLKNAVKE 181
DB 122 ksehavlgldnltdnagrailkkhankdkgaalekrlfkavenlskaadtlknave 181
OY 181 LTPSVIHGNNRKGDNASTNSADESVKGPNTLTKITSSNAVYLAKEVETLLASIDE 240
DB 181 ltpsvihgnnsrkdgnastnsadesvkgpnlteiskkltesnavylavkevetllaside 240
OY 241 ATKAIGKIGKNGNGLEANSOKNTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAACOSTE 301
DB 241 atkaigkigngngleangskntsllsgayaissdlaeklnvlkneelkekidtakgscst 301
OY 301 EFTNKLKSEHAVLGLDNLTDNAORAILKKHANKDKGALEKLFKAVENTLSKAADTLK 361
DB 301 eftnklksehavlgldnltdnagrailkkhankdkgaalekrlfkavenlskaadtlkn 361
OY 362 AVKELTS 368
DB 362 avkeltss 368

RESULT 3

AAB62719
ID AAB62719 standard; Protein: 367 AA.

XX AAB62719;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 42.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN W0200078966-A1.

PD 28-DEC-2000.

PE 19-JUN-2000; 2000WO-US16915.

PI 18-JUN-1999; 99US-0140042.

```

XX (UUNY ) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
PI Dattwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX MPI: 2001-050113/06.
DR N-PSDB: AAF29023.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43: Page 102-103; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 367 AA;
SQ
XX
XX Query Match 85.3%; Score 1533; DB 22; Length 367;
XX Best Local Similarity 86.6%; Pred. No. 1.1e-90;
XX Matches 318; Conservative 20; Mismatches 27; Indels 2; Gaps 2;
XX
XX 2 ACNNSGKDGNASANSADSVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDELATKAI 61
XX 1 acnngskdgnasansadesvkgpnlteiskitesnavlavkevellasidelatka 60
XX
XX 62 GKRTGNGGLEANSKNTSLSGAYAISDLIAEKLNVLNKNELEKIDTAKOCSTFTNKL 121
XX 61 gkrtgnggleangskntsllsgayaaisdliaeklnvlnknelekidtakocstftnkl 120
XX
XX 122 KSHAVGLDNLTDNNAQRALIKKHANKDKGALEKLFKAVENLSKRAODTLKNAYKEL 181
XX 121 kshavglldnltddnagrallkxhankdkgaalekelfkavenlskaadclknaykel 180
XX
XX 182 TSPVHGNNRSKRDGNASTNSADESVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDEL 241
XX 181 tspvhnngskrdgntnsadesvkgpnlteiskitesnavlavkeletllasidel 240
XX
XX 242 ATTAIGKKT-GNNGLEANSKNTSLSGAYAISDLIAEKLNVLNK-BELKEKIDTAKOC 299
XX 241 attaigkktgngnglaveagngllagayctlsklltqkldglnsekliketienakcs 300
XX
XX 300 TEFNTNKLKSHAVGLDNLTDNNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTL 359
XX 301 tefntnklkshavglldnltddnagrallkxhankdkgaalekelfkavenlskaadclknaykel 360
XX
XX 360 KNAVVEL 366
XX 361 ansvkel 367
XX
XX
XX RESULT 4
XX AAB62735
XX ID AAB62735 standard; Protein: 391 AA.
XX
XX AAB62735:
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 74.
XX
XX Borrelia: ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
XX
XX Chimeric - Borrelia sp.
XX
XX WO200078966-A1.

```

```

XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000WO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Dattwyler RJ, Seinoost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX MPI: 2001-050113/06.
XX DR N-PSDB: AAF29039.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43: Page 142; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 391 AA;
SQ
XX
XX Query Match 85.1%; Score 1530; DB 22; Length 391;
XX Best Local Similarity 86.4%; Pred. No. 1.8e-90;
XX Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;
XX
XX 2 ACNNSGKDGNASANSADSVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDELATKAI 61
XX 25 acnngskdgnasansadesvkgpnlteiskitesnavlavkevellasidelatka 84
XX
XX 62 GKRTGNGGLEANSKNTSLSGAYAISDLIAEKLNVLNKNELEKIDTAKOCSTFTNKL 121
XX 85 gkrtgnggleangskntsllsgayaaisdliaeklnvlnknelekidtakocstftnkl 144
XX
XX 122 KSHAVGLDNLTDNNAQRALIKKHANKDKGALEKLFKAVENLSKRAODTLKNAYKEL 181
XX 145 kshavglldnltddnagrallkxhankdkgaalekelfkavenlskaadclknaykel 204
XX
XX 182 TSPVHGNNRSKRDGNASTNSADESVKGPNTLEISKITTESNAVAVLAVKEVETLLASIDEL 241
XX 205 tspvhnngskrdgntnsadesvkgpnlteiskitesnavlavkeletllasidel 264
XX
XX 242 ATTAIGKKT-GNNGLEANSKNTSLSGAYAISDLIAEKLNVLNK-BELKEKIDTAKOC 299
XX 265 attaigkktgngnglaveagngllagayctlsklltqkldglnsekliketienakcs 324
XX
XX 300 TEFNTNKLKSHAVGLDNLTDNNAQRALIKKHANKDKGALEKLFKAVENLSKRAADTL 359
XX 325 tefntnklkshavglldnltddnagrallkxhankdkgaalekelfkavenlskaadclknaykel 384
XX
XX 360 KNAVVEL 366
XX 385 ansvkel 391
XX
XX
XX RESULT 5
XX AAB62718
XX ID AAB62718 standard; Protein: 368 AA.
XX
XX AAB62718:
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 40.

```

KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 XX Chimeric - Borrelia sp.
 PN WO20078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UNYNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI; 2001-050113/06.
 DR N-PSDB; AAF29022.
 XX
 PT Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 PS Claim 43; Page 99-100; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospC proteins and
 CC chimeric ospC proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. atzei or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 368 AA;

Query Match 82.7%; Score 1485.5; DB 22; Length 368;
 Best Local Similarity 84.3%; Pred. No. 1.2e-87;
 Matches 311; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNNGSGKGNASANSADSVKGNPLTEISKRTTESNAVLAVKEVETLLASIDELATKAI 61
 DB 1 acnngsgkgnasansadesvkgpnlteliskrttesnavlavkevetllasidelatkai 60
 QY 62 GKRIKNGGLEANSQKNTSLSGAYAISDLIAEKLNVLENELEKIDTAKOCSTEFNTKL 121
 DB 61 gkriknnggleansqkntslsgayaidsdliaeklnvleekidetakgstefntkl 120
 QY 122 KSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAOITLKNAYVEL 181
 DB 121 ksehavlgldnltdnnaoraillkhanndkgaalekfkfavenstkaaqdtlknayvel 180
 QY 182 TSPYHGNNSRKDGNASTNSADESVKGNPLTEISKRTTESNAVLAVKEVETLLASIDEL 241
 DB 181 tsplhgnnsgkdnstnsadesvkgpnlteliskrttesnavlavkevetllasidel 240
 QY 242 ATKAIGKRIGNN-GLEANQKNTSLSGAYAISDLIAEKLNVLENE-ELKEKIDTAKOCS 299
 DB 241 a-kaigkriknng-gleanqkntslsgayaidsdliaeklnvleekidetakgstefntkl 299
 QY 300 TEFTNKLKSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAOITL 359
 DB 300 teftnklksehavlgldnltdnnaoraillkhanndkgaalekfkfavenstkaaqdtlknayvel 359
 QY 360 KNAVELTSS 368
 DB 360 tnsvelts 368

RESULT 6
 AAB62734
 ID AAB62734 standard; Protein: 392 AA.
 XX

AC AAB62734;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospC protein SEQ ID NO: 72.
 XX
 KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 XX Chimeric - Borrelia sp.
 PN WO20078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UNYNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 DR WPI; 2001-050113/06.
 DR N-PSDB; AAF29038.
 XX
 PT Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 PS Claim 43; Page 139-140; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospC proteins and
 CC chimeric ospC proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. atzei or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 392 AA;

Query Match 82.5%; Score 1482.5; DB 22; Length 392;
 Best Local Similarity 84.0%; Pred. No. 2e-87;
 Matches 310; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNNGSGKGNASANSADSVKGNPLTEISKRTTESNAVLAVKEVETLLASIDELATKAI 61
 DB 25 acnngsgkgnasansadesvkgpnlteliskrttesnavlavkevetllasidelatkai 84
 QY 62 GKRIKNGGLEANSQKNTSLSGAYAISDLIAEKLNVLENELEKIDTAKOCSTEFNTKL 121
 DB 85 gkriknnggleansqkntslsgayaidsdliaeklnvleekidetakgstefntkl 144
 QY 122 KSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAOITLKNAYVEL 181
 DB 145 ksehavlgldnltdnnaoraillkhanndkgaalekfkfavenstkaaqdtlknayvel 204
 QY 182 TSPYHGNNSRKDGNASTNSADESVKGNPLTEISKRTTESNAVLAVKEVETLLASIDEL 241
 DB 205 tsplhgnnsgkdnstnsadesvkgpnlteliskrttesnavlavkevetllasidel 264
 QY 242 ATKAIGKRIGNN-GLEANQKNTSLSGAYAISDLIAEKLNVLENE-ELKEKIDTAKOCS 299
 DB 265 a-kaigkriknng-gleanqkntslsgayaidsdliaeklnvleekidetakgstefntkl 323
 QY 300 TEFTNKLKSEHAVLGLDNLTDNNAORAILKKHANKDKGAELFKFKAVENTSKAAOITL 359
 DB 324 teftnklksehavlgldnltdnnaoraillkhanndkgaalekfkfavenstkaaqdtlknayvel 383
 QY 360 KNAVELTSS 368
 DB 384 tnsvelts 392

```

RESULT 7
AAB62714
ID AAB62714 standard; Protein; 369 AA.
XX
AC AAB62714;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 32.
XX
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX
PI (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
XX
DR N-PSDB: AAF29018.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
XX
PS Claim 43; Page 89; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
XX
CC chimeric ospC proteins from members of the Borrelia genus. These may be
XX
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX
CC vaccines against Borrelia infection, which is spread by ticks and leads
XX
CC to Lyme disease.
XX
SQ Sequence 369 AA;

Query Match 82.3%; Score 1479; DB 22; Length 369;
Best Local Similarity 83.8%; Pred. No. 3e-87;
Matches 310; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

```

```

DB 300 stfntklskshavlgldntldnnaqraillkhanckgaalekltkavenlskaagdt 359
QY 359 LKNAVKELTG 368
DB 360 lknavkeltg 369

RESULT 8
AAB62732
ID AAB62732 standard; Protein; 393 AA.
XX
AC AAB62732;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 68.
XX
KM Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX
PI (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06.
XX
DR N-PSDB: AAF29036.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
XX
PT Lyme disease are used to immunize animals and detect immune responses
XX
PT to Lyme disease -
XX
PS Claim 43; Page 133-134; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
XX
CC chimeric ospC proteins from members of the Borrelia genus. These may be
XX
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX
CC vaccines against Borrelia infection, which is spread by ticks and leads
XX
CC to Lyme disease.
XX
SQ Sequence 393 AA;

Query Match 82.1%; Score 1476; DB 22; Length 393;
Best Local Similarity 83.5%; Pred. No. 5.1e-87;
Matches 309; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

```


QY 62 GKRI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVLNKELEKIDTAOKOSTEFTNK 120
 DB 61 gkrlhngldteyhnngslagayalstlkqldglneglkexidaakkcssetfntk 120
 QY 121 LKSEHAVLGIDNTDDNAORAILKHKANKDGAAELEKLFKAVENTSKAAOTLKNVKE 180
 DB 121 lkehtldgkyvtdadakaellktngctktgkeelgkfesevytskaakemlansvke 180
 QY 181 LTPSPYHG-----NSNRKDGNASTNSADESVKGNPNTLEISKRTITESNAVAVLAVEVETLL 235
 DB 181 ltspvvaespmwvngskdgtansadesvkgpnlteiskrtitesnavlavkeletll 240
 QY 236 ASIDELATATKIKKI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVLNKN-EELKEKID 293
 DB 241 asidelatkaigkllqngslaveagngllagaylisklltqkldglnseklike 300
 QY 294 TAKOCSTEFNTKLSEHAVLGIDNTDDNAORAILKHKANKDGAAELEKLFKAVENTLSK 353
 DB 301 nakccsedftkklegahqiglenvdenakkailltdaakdgaaeleklfkavenlak 360
 QY 354 AAOTLKNVAVKEL 366
 DB 361 aakemlansvkel 373
 RESULT 11
 AAB62729
 ID AAB62729 standard; Protein; 397 AA.
 AC AAB62729;
 XX
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 62.
 XX
 KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000MO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UYNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Seinoest G, Dykhuzen D, Luft BJ, Gomes-Solecki M;
 DR MPI: 2001-050113/06.
 DR N-PSDB: AAF29033.
 XX
 PT Compositions of Ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 PS Claim 43; Page 125-126; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 397 AA;
 Query Match 67.9%; Score 1221; DB 22; Length 397;

Best Local Similarity 70.0%; Pred. No. 1e-70;
 Matches 261; Conservative 43; Mismatches 61; Indels 8; Gaps 4;
 QY 2 ACNNSKDGNAANSADSVKGNPNTLEISKRTITESNAVAVLAVEVETLLASIDELATKAI 61
 DB 25 scnsgskdgtntsansadesvkgpnlteiskrtitesnavlavkeveallssideiaakal 84
 QY 62 GKRI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVLNKELEKIDTAOKOSTEFTNK 120
 DB 85 gkrlhngldteyhnngslagayalstlkqldglneglkexidaakkcssetfntk 144
 QY 121 LKSEHAVLGIDNTDDNAORAILKHKANKDGAAELEKLFKAVENTSKAAOTLKNVKE 180
 DB 145 lkehtldgkyvtdadakaellktngctktgkeelgkfesevytskaakemlansvke 204
 QY 181 LTPSPYHG-----NSNRKDGNASTNSADESVKGNPNTLEISKRTITESNAVAVLAVEVETLL 235
 DB 205 ltspvvaespmwvngskdgtansadesvkgpnlteiskrtitesnavlavkeletll 264
 QY 236 ASIDELATATKIKKI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVLNKN-EELKEKID 293
 DB 265 asidelatkaigkllqngslaveagngllagaylisklltqkldglnseklike 324
 QY 294 TAKOCSTEFNTKLSEHAVLGIDNTDDNAORAILKHKANKDGAAELEKLFKAVENTLSK 353
 DB 325 nakccsedftkklegahqiglenvdenakkailltdaakdgaaeleklfkavenlak 384
 QY 354 AAOTLKNVAVKEL 366
 DB 385 aakemlansvkel 397
 RESULT 12
 AAB62716
 ID AAB62716 standard; Protein; 369 AA.
 XX
 AC AAB62716;
 XX
 DT 03-APR-2001 (first entry)
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 36.
 XX
 KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000MO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX
 PA (UYNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Datwyler RJ, Seinoest G, Dykhuzen D, Luft BJ, Gomes-Solecki M;
 DR MPI: 2001-050113/06.
 DR N-PSDB: AAF29020.
 XX
 PT Compositions of Ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 PS Claim 43; Page 94-95; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads

CC to Lyme disease.

XX Sequence 369 AA;

Query Match 67.9%; Score 1220; DB 22; Length 369;
 Best Local Similarity 70.5%; Pred. No. 1.1e-70;
 Matches 261; Conservative 45; Mismatches 58; Indels 6; Gaps 6;

QY 2 ACNNSGKDNASANSADSVKGNLTETSKRTTESNAVLAVKEVEETLLASIDELATKAI 61
 DB 1 acnnsygdntsansadesvkgpnlteiskrtidsnavllavkeveallssidela-kai 59
 QY 62 GKRTGNNG-LEFANOSKNTSLSGAVALSDLAELKLVKNEE-LKEKIDTAKOCSTEF 119
 DB 60 gkklngdgsldneanresllagaystltltqklsklingseggkelaakcseefst 119
 QY 120 KLKSEHVLGLDNLTDNAORALIKKH-ANKDGALEKLFRAVENLSKAADTLKNAY 178
 DB 120 klkshhaglgvgtldenakkaillkanaagkdkyveeleklsglseslskaakemiansv 179
 QY 179 KELTSPVHGNNSRKDNASTNSADESVKGNLTETSKRTTESNAVLAVKEVEETLLAS 238
 DB 180 keltspvvhgnnsygdntsansadesvkgpnlteiskrtidsnavllavkeveetllasi 239
 QY 239 DELATKAIGKRT-GNNGLEFANOSKNTSLSGAVALSDLAELKLVKN-EELKEKIDTAK 296
 DB 240 delatkaigkrtgngglaveagngtllagaytiskltqkldgklnseklkelenak 299
 QY 297 QCSTEFNTKLSKSEHAVLGLDNLTDNAORALIKKHANKDKGALEKLFRAVENLSKAQ 356
 DB 300 kcsedftkklgeghaglgienvndenakkailltdaakdkgaaleklfkavenlakaak 359
 QY 357 DTLKNAVVEL 366
 DB 360 emlansvkel 369

RESULT 13

AAB62739

ID AAB62739 standard; Protein; 400 AA.

AC AAB62739;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

XX 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

PR (UYNV) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

XX N-PSDB; AAF29043.

XX Compositions of ospc polypeptides from strains of Borrelia which cause

XX Lyme disease are used to immunize animals and detect immune responses

XX to Lyme disease -

PS Claim 43; Page 153; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

SQ Sequence 400 AA;

Query Match 67.9%; Score 1219.5; DB 22; Length 400;
 Best Local Similarity 69.4%; Pred. No. 1.3e-70;
 Matches 261; Conservative 43; Mismatches 61; Indels 11; Gaps 4;

QY 2 ACNNSGKDNASANSADSVKGNLTETSKRTTESNAVLAVKEVEETLLASIDELATKAI 61
 DB 25 scnnsygdntsansadesvkgpnlteiskrtidsnavllavkeveallssidelaakai 84
 QY 62 GKRT-GNNGLEFANOSKNTSLSGAVALSDLAELKLVKNEE-LKEKIDTAKOCSTEF 120
 DB 60 gkklngdgsldneanresllagaystltltqklsklingseggkelaakcseefst 144
 QY 121 IKSEHAVLGLDNLTDNAORALIKKHANKDKGALEKLFRAVENLSKAADTLKNAY 180
 DB 145 lkehtldgkegytdadekaalktngtktygaeeigkllsesveylskaakemiansvke 204
 QY 181 LTSPVHNG-----NNSRKDNASTNSADESVKGNLTETSKRTTESNAVLAVKEVE 232
 DB 205 ltpsvaesppkpsvnsygdntsansadesvkgpnlteiskrtidsnavllavkele 264
 QY 233 TLASIDELATKAIGKRT-GNNGLEFANOSKNTSLSGAVALSDLAELKLVKN-EELKE 290
 DB 265 tlasidelaakagkkgngglaveagngtllagaytiskltqkldgklnseklk 324
 QY 291 KIDTAKOCSTEFNTKLSKSEHAVLGLDNLTDNAORALIKKHANKDKGALEKLFRAVEN 350
 DB 325 klenakcsedftkklgeghaglgienvndenakkailltdaakdkgaaleklfkaven 384
 QY 351 LSKAADTLKNAYVEL 366
 DB 385 lakaakemiansvkel 400

RESULT 14

AAB62731

ID AAB62731 standard; Protein; 393 AA.

AC AAB62731;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 66.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

XX Chimeric - Borrelia sp.

PN WO200078966-A1.

XX 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

PR (UYNV) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29035.
 XX Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 PS Claim 43; Page 131; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 393 AA;
 Query Match 67.7%; Score 1217; DB 22; Length 393;
 Best Local Similarity 70.3%; Pred. No. 1.8e-70;
 Matches 260; Conservative 46; Mismatches 58; Indels 6; Gaps 6;
 QY 2 ACNNSGKDNASNSADESVKGNLTETSKITESNAVLAKEVEETLLASIDELATKAI 61
 DB 25 scngskdgttsansadesvkgpnlteiskitdsnavllavreveallsidele-kai 83
 QY 62 GKRIKNGG-LEANOSKNTSLSGAIVASIDIAEKLVNKNKE-LKEKIDTAKOCSTETN 119
 DB 84 gkrlknggsldeanrnesllagaytstltqtklsgseglkeklaakcseefst 143
 QY 120 KLTSEHAVGLDNLTDNAORAILKKH-ANKDKGAELKLFKAVENLSKAODTLKNAY 178
 DB 144 kltshavglgldntdenakalkkanaagkdyveelksgleslskaakemlansv 203
 QY 179 KELTSPIVHGNNSRKDNASTNSADESVKGNLTETSKITESNAVLAKEVEETLLASI 238
 DB 204 keltspvhnsgkdgttsansadesvkgpnlteiskitteenavllavkeletlasi 263
 QY 239 DELATAIKGKTI-GNNGLEANSKNTSLSGAIVASIDIAEKLVNKN-ELKKEIDTAK 296
 DB 264 delatalkgkllqngslaveaghnllagaytlslkldglnseklkeklenak 323
 QY 297 QCSSTEFNKLKSEHAVGLDNLTDNAORAILKKHANKDKGAELKLFKAVENLSKRAO 356
 DB 324 kcsedftkklgeghagdytgvdenakallltidaekdkgaeeleklfkavenlakaak 383
 QY 357 DTLKNAVKEL 366
 DB 384 emlansvkel 393
 RESULT 15
 AAB62712
 ID AAB62712 standard; Protein; 378 AA.
 XX
 AC AAB62712;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE Borrelia sp chimeric ospc protein SEQ ID NO: 28.
 XX
 KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
 XX
 OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.
 PN WO200078966-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000WO-US16915.
 XX
 PR 18-JUN-1999; 99US-0140042.
 XX

PA (UNY) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 XX
 PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 XX
 DR WPI: 2001-050113/06.
 DR N-PSDB; AAF29016.
 XX
 CC Compositions of ospc polypeptides from strains of Borrelia which cause
 CC Lyme disease are used to immunize animals and detect immune responses
 CC to Lyme disease -
 XX
 PS Claim 43; Page 83-84; 160pp; English.
 XX
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX
 SQ Sequence 378 AA;
 Query Match 65.6%; Score 1178; DB 22; Length 378;
 Best Local Similarity 67.3%; Pred. No. 5.3e-68;
 Matches 235; Conservative 47; Mismatches 65; Indels 12; Gaps 5;
 QY 1 MACNNSGKDNASNSADESVKGNLTETSKITESNAVLAKEVEETLLASIDELATKA 60
 DB 1 macnsgkdgttsansadesvkgpnlteinkitdsnavllavreveallsideleaka 60
 QY 61 IGKTI-GNNGLEANSKNTSLSGAIVASIDIAEKLVNKNKEELKEKIDTAKOCSTETN 119
 DB 61 igkthngldtennngslagaystltqtklsgseglkekidaekcseftn 120
 QY 120 KLTSEHAVGLDNLTDNAORAILKKHANKDKGAELKLFKAVENLSKAODTLKNAY 179
 DB 121 kltshavglgldntdenakalkkanaagkdyveelksgleslskaakemlansv 180
 QY 180 ELTSPIVHG-----NSRKDNASTNSADESVKGNLTETSKITESNAVLAKEV 231
 DB 181 eltpvvaespkkpmsvnsngkdgttsansadesvkgpnlteiskitteenavllavkev 240
 QY 232 ETLASIDELATKAIGKRIKNGG-LEANOSKNTSLSGAIVASIDIAEKLVNKN-ELK 289
 DB 241 etllsidele-kai gkrlknggsldeanrnesllagaytstltqtklsgseglkek 299
 QY 290 EKIDTAKOCSTEFNKLKSEHAVGLDNLTDNAORAILKKHANKDKGAELKLFKAVE 349
 DB 300 aeiekakcseefcaklgehtdgykgvtdnnakalkkcncktkgadelkelfesvk 359
 QY 350 NLSKAODTLKNAYKELTS 368
 DB 360 nlskaakemlansvkelts 378

Search completed: March 18, 2002, 09:54:33
 Job time: 331 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:34 ; Search time 55.5 Seconds

(without alignments)
149.211 Million cell updates/sec

Title: US-09-596-746A-38

Perfect score: 1797

Sequence: 1 MACNNSGKDGDNASANSADSE.....ENLSKAQDTLKNVKELTLS 368

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA: *

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
- 5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep: *
- 6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	37.5	466	4	US-08-235-836C-110 Sequence 110, App
2	657.5	36.6	212	1	US-08-158-353-2 Sequence 2, App11
3	640	35.6	588	4	US-08-235-836C-122 Sequence 122, App
4	638.5	35.5	466	4	US-08-235-836C-107 Sequence 107, App
5	623.5	34.7	209	4	US-09-196-293-15 Sequence 15, App1
6	623.5	34.7	210	1	US-08-158-353-3 Sequence 3, App1
7	617.5	34.4	210	4	US-08-235-836C-30 Sequence 30, App1
8	616.5	34.3	210	4	US-08-209-603E-15 Sequence 15, App1
9	588.5	32.7	212	1	US-08-158-353-4 Sequence 4, App1
10	584.5	32.5	212	1	US-08-031-295-2 Sequence 2, App1
11	584.5	32.5	212	4	US-07-903-580-2 Sequence 2, App1
12	583.5	32.5	212	4	US-08-235-836C-34 Sequence 34, App1
13	578.5	32.2	212	4	US-09-196-293-11 Sequence 11, App1
14	578.5	32.2	212	4	US-08-209-603E-11 Sequence 11, App1
15	563	31.3	207	4	US-08-235-836C-36 Sequence 36, App1
16	536	29.8	209	4	US-08-235-836C-32 Sequence 32, App1
17	157.5	8.8	1588	5	PCT-US93-07261-11 Sequence 11, App1
18	157.5	8.8	1663	5	PCT-US93-07261-16 Sequence 16, App1
19	150	8.3	695	1	US-08-127-499A-23 Sequence 23, App1
20	150	8.3	695	1	US-08-482-847-23 Sequence 23, App1
21	150	8.3	1147	3	US-08-470-260-5 Sequence 5, App1
22	150	8.3	1147	3	US-08-471-491-5 Sequence 5, App1
23	150	8.3	1147	4	US-08-466-662-5 Sequence 5, App1
24	150	8.3	3289	2	US-08-477-451-2 Sequence 2, App1
25	149.5	8.3	1196	1	US-08-144-121-4 Sequence 4, App1
26	149.5	8.3	1196	2	US-08-735-893-4 Sequence 4, App1
27	149	8.3	630	4	US-08-973-462-9 Sequence 9, App1

28	141.5	7.9	1786	4	US-08-973-462-8 Sequence 8, App1
29	141	7.8	859	1	US-08-053-614-2 Sequence 2, App1
30	141	7.8	859	1	US-08-316-397B-2 Sequence 2, App1
31	141	7.8	859	2	US-09-034-306-2 Sequence 2, App1
32	141	7.8	859	4	US-09-259-437-2 Sequence 2, App1
33	141	7.8	859	5	PCT-US93-09782-2 Sequence 2, App1
34	141	7.8	1181	1	US-08-053-614-4 Sequence 4, App1
35	141	7.8	1181	1	US-08-316-397B-4 Sequence 4, App1
36	141	7.8	1181	2	US-09-034-306-4 Sequence 4, App1
37	141	7.8	1181	5	PCT-US93-09782-4 Sequence 4, App1
38	141	7.8	1181	5	PCT-US93-09782-4 Sequence 4, App1
39	139.5	7.7	2482	1	US-08-328-254-6 Sequence 6, App1
40	138.5	7.7	619	1	US-08-465-746-2 Sequence 2, App1
41	138.5	7.7	619	1	US-08-214-164-2 Sequence 2, App1
42	138.5	7.7	619	2	US-08-467-852A-3 Sequence 3, App1
43	138.5	7.7	619	2	US-08-246-636-2 Sequence 2, App1
44	138.5	7.7	619	2	US-08-247-491A-3 Sequence 3, App1
45	138.5	7.7	619	2	US-08-319-795-2 Sequence 2, App1

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562zel Chimeric Proteins Comprising
TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESS: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235, 836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110

Query Match 37.5%; Score 673; DB 4; Length 466;
Best Local Similarity 44.9%; Pred. No. 1.3e-45;
Matches 173; Conservative 58; Mismatches 112; Indels 42; Gaps 11;
OY 1 MACNNSGKDGDNASANSADSEYKGNPLTEISKKTIESNAYVLAVKEVETLLASIDELATKA 60

```
Db 17 ISCNNSGKDGNTSANSADSVKGNLTETISKTIIDSNVAVLAVKEVETLLASIDELATAKA 76
OY 61 IGGKI-GANGLEANOSKNTSLSGAVATSDLIAEKLNVAKNELEKIDTAACOSTEFTN 119
Db 77 IGGKIHONGNGIDTETYNHNSGLARVAISTLIKOKIDGLKNEGLEKIDAKKCEFTFN 136
OY 120 KIKSEHAVIGLNDLTJDNQORAILKHKHANKDGAELKFKAVENLSKAADOTLKNAVK 179
Db 137 KIKKHHIDLGEGVTDADAKAIIKTGTGTGAEELKLFESVEVLAKAKELNLSVK 196
OY 180 ELTSPPIYHGNNSRKDGNAST---NSADESVKGNLTETISKTIITSNVAVLAVKEVETL 235
Db 197 ELTSPVVAESPKPKQVSSVLEKNSVVDLPGEKVLVSKENKCKOY-----DLI 248
OY 236 ASIDELATKAIIGKIGNN-----GLEANSKNTSLSGAVATSDLIAEKLNVAKNEE-- 287
Db 249 ATYDKLELKTSDK--NNGSGVLEGVAKDSKVKLTIS-----DDIGQTTLEVFKEKDGKT 301
OY 288 -LKEKIDTAACOSTE--FTNKLKSEHAVIGLNDLTJDNQORAILKHKHANKDGAELKFK 344
OY 302 LVSKKVTSKDSSTEEKFNEKEVSEKII-----TRADGTLEVTGTGKSDSGKAKEY 354
OY 345 FK--AVENLSKAADOTLKNAVKELT 367
Db 355 LKGYVLEGTILTAERTTL--VVKEGT 377

RESULT 2
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UCT93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9340
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-158-353-2

Query Match 36.6%; Score 657.5; DB 1; Length 212;
Best Local Similarity 72.9%; Pred. No. 7,6e-45;
Matches 140; Conservative 20; Mismatches 29; Indels 3; Gaps 2;
```

```
OY 1 MACNNGKDGNAANSADSVKGNLTETISKTIITSNVAVLAVKEVETLLASIDELATAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNLTETISKTIITSNVAVLAVKEVETLLASIDELATAKA 76
OY 61 IGGKIHONGNGIDTETYNHNSGLARVAISTLIKOKIDGLKNEGLEKIDAKKCEFTFN 119
Db 77 IGGKIHONGNGIDTETYNHNSGLARVAISTLIKOKIDGLKNEGLEKIDAKKCEFTFN 136
OY 120 KIKSEHAVIGLNDLTJDNQORAILKHKHANKDGAELKFKAVENLSKAADOTLKNAVK 177
Db 137 KIKSSAEELGINGAATTDANAKAAILKTGTGKAELEKLFESVKMLSKRAADOTLNN 196
OY 178 VKELTSPPIYHGN 189
Db 197 VKELTSPVVAEN 208

RESULT 3
US-08-235-836C-122
; Sequence 122, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESS: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-Apr-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNI93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-836C-122

Query Match 35.6%; Score 640; DB 4; Length 588;
Best Local Similarity 39.1%; Pred. No. 6,8e-43;
Matches 184; Conservative 55; Mismatches 110; Indels 122; Gaps 17;
```

```

Db 69 QVELKGTGSDK--NNGSGTLEGSKPKSKYKLVASADLNTVLEAFDA-SNQKISSKV-TK 124
Qy 111 KQCS-TEPT---NKL-----KSEHAVLGLDNLTD-DNORAI-----142
Db 125 KQSITEETLKAKKLSKRLTRNSGTTLEYSQITDADNATKAVETLKNISIKLEGLVVGK 184
Qy 143 -----LKKHANKD-----150
Db 185 TTYVEIKEGVTLTKREIEKDKQVAVFLNDTFRAGSKKTKGKEDSTSTLTISADSKTKDVLV 244
Qy 151 -----KGAELEKLFKRAVENLSKRAODTLKNAVELSPVTHGNSRKDNAS 198
Db 245 LTDCGTTVOOYNAGTSLESASEIKNLSE-----LKNALK--GHPH--GNNSGKQNTS 295
Qy 199 TNSADESVKGNLTETISKTTESNAVAVLAVKEVETLTIASIDELATKAIKKTI-GNNGLEA 257
Db 296 ANSADESVPKPNLTETISKTTDSNAVAVLAVKEVEALLSIDETIAAKIGKKIHQNNGLDT 355
Qy 258 NOSKNISLSGAAVAISDLIAEKLNVLKNELKEKIDTAKCSTETPKKSEHAVLGLDN 317
Db 356 EYHNNSLLAGRYAISTLTKOKLDGKNEGKLEKIDAKKCSFTPKKLEKHTDLGKEG 415
Qy 318 LTDFDNORAILKKHANKDGALEKLFKRAVENLSKRAODTLKNAVELTS 368
Db 416 VTDADAKKALTKNGTKTKGAELGKLFESVEVLSKAKEMLANSVKELTS 466

RESULT 4
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESS: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-107
Query Match 35.5%; Score 638.5; DB 4; Length 466;

```

```

Best Local Similarity 44.6%; Pred. No. 6,6e-43;
Matches 166; Conservative 50; Mismatches 119; Indels 37; Gaps 8;

Qy 19 ESKVGNLTETISKTTESNAVAVLAVKEVETLTIASIDELATKAIKKTI-GNNGLEA-SK 77
Db 100 EFKEDKGTLSKVTSKDSSTEEKFNEKEVS-EKIIIRADGTRELYGIDKSDGSGKA 158
Qy 78 TSLISGAAVAISDLIAEKLNVLKNELKEKIDTAKCSTETPKKSEHAVLGLDNLT 134
Db 159 KEVLKGYVLEGTTLAEKTLVKEGTVTLKRNLSKSGSEVELNDTSS--AATKTRAW 216
Qy 135 DDNAORAILKKHANKD-----KGAELEKLFKRAVENLSKRAODTLKNA 177
Db 217 NDSTSTLTISADSKTKTDVFLTDCGTTVOOYNAGTSLESASEIKNLSE-----LKNA 271
Qy 178 VKELTSPVTHGNSRKDNASTNSADESVKGNLTETISKTTESNAVAVLAVKEVETLTIAS 237
Db 272 LK-----MANNSGKDNSTNSADESVKGNLTETISKTTDSNAVAVLAVKEVEALLSS 324
Qy 238 IDELATKAIKKTI-GNNGLEANSKNTSLSGAAVAISDLIAEKLNVLKNELKEKIDTAK 296
Db 325 IDEIAARAIGKKIHQNNGLDTVEYHNNSLLAGRYAISTLTKOKLDGKNEGKLEKIDAK 384
Qy 297 QCSFTETPKKSEHAVLGLDNLTDNORAILKKHANKDGALEKLFKRAVENLSKRAO 356
Db 385 KCSETFTNKLAKHTDLCGEGVTDADAKKALTKNGTKTKGAELGKLFESVEVLSKAK 444
Qy 357 DTLKNAVELTS 368
Db 445 EMLANSVKELTS 456

RESULT 5
US-09-196-293-15
; Sequence 15, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; APPLICANT: Fuchs, Renate
; APPLICANT: Motz, Manfred
; APPLICANT: Sautscheck, Erwin
; APPLICANT: Wilske, Bettina
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrella
; FILE REFERENCE: 738.001052
; CURRENT APPLICATION NUMBER: US/09/196,293
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1989-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; EARLIER FILING DATE: 1990-06-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Borrella burgdorferi
US-09-196-293-15
Query Match 34.7%; Score 623.5; DB 4; Length 209;
Best Local Similarity 69.3%; Pred. No. 3.5e-42;
Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;
Qy 1 MACNNSGKDGNAANSADSVKGNLTETISKTTESNAVAVLAVKEVETLTIASIDELATKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNLTETISKTTDSNAVAVLAVKEVEALLSIDETIAKA 76

```

```

QY      61   IGGKI-GNNGLAANOSKRTSLILSGAVAIISDILAERKLNVLKNBELEKEKIDTAKOCSTEFNTN 119
        |||||  ||||| : | |||:||||| | :||: ||| ||||| ||: || |||
Db      77   IGGKIHHONGIGDTENNHNHNSLLAGAVIAISTLIKOKLIDGKLNGLKEKIDAAKKCSEFTFN 136
        |||||  ||||| : | |||:||||| | :||: ||| ||||| ||: || |||
QY      120  KUKSEHAVLIGDNLTDJDNACRAIIRKKNANKDKGAELFEKIFRAVENTSKAODTLKNAVK 179
        ||| : | : || : || : ||| : | ||| || ||||| : ||| ||| : |||
Db      137  KUKSEHTDLDGEGVYADDAKAFAILKTNGTKTGAEELGKLFESVEVLSSKAKEMLANSVK 196
        ||| : | : || : || : ||| : | ||| || ||||| : ||| ||| : |||
QY      180  ELTSPIV 186
        ||||| : |
Db      197  ELTSPVV 203
        ||||| : |

RESULT      6
; Sequence 3, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UC793-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEO ID NO.: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-158-353-3

Query Match          34.7%; Score 623.5; DB 1; Length 210;
Best Local Similarity 69.5%; Pred. No. 3,6e-42;
Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1

QY      1 MACNNCGKGCNASANSADSVGVGNPTIELSKITTESNAVYLAVEKETLLASIDELATYA 60
        |||||  ||||| : | |||:||||| | :||: ||| ||||| ||: || |||
Db      17   ISCCNNGKRGKNTSANSADSVGVGNPTIELSKITTSNAVLLVAKEVALLSIDEATAA 76
        |||||  ||||| : | |||:||||| | :||: ||| ||||| ||: || |||
QY      61   IGKTI-GNNGLAANOSKRTSLILSGAVAISDLIAEKLVNKNELEKEKIDTAOCSTEFNTN 119
        |||||  ||||| : | |||:||||| | :||: ||| ||||| ||: || |||
Db      77   IGGKIHHONGIGDTENNHNHNSLLAGAVIAISTLIKOKLIDGKLNGLKEKIDAAKKCSEFTFN 136
        |||||  ||||| : | |||:||||| | :||: ||| ||||| ||: || |||
QY      120  KUKSEHAVLIGDNLTDJDNACRAIIRKKNANKDKGAELFEKIFRAVENTSKAODTLKNAVK 179
        ||| : | : || : || : ||| : | ||| || ||||| : ||| ||| : |||
Db      137  KUKSEHTDLDGEGVYADDAKAFAILKTNGTKTGAEELGKLFESVEVLSSKAKEMLANSVK 196
        ||| : | : || : || : ||| : | ||| || ||||| : ||| ||| : |||
QY      180  ELTSPIV 186
        ||||| : |
Db      197  ELTSPVV 203
        ||||| : |

```

```

      1
      2
      3
      4
      5
      6
      7
      8
      9
     10
     11
     12
     13
     14
     15
     16
     17
     18
     19
     20
     21
     22
     23
     24
     25
     26
     27
     28
     29
     30
     31
     32
     33
     34
     35
     36
     37
     38
     39
     40
     41
     42
     43
     44
     45
     46
     47
     48
     49
     50
     51
     52
     53
     54
     55
     56
     57
     58
     59
     60
     61
     62
     63
     64
     65
     66
     67
     68
     69
     70
     71
     72
     73
     74
     75
     76
     77
     78
     79
     80
     81
     82
     83
     84
     85
     86
     87
     88
     89
     90
     91
     92
     93
     94
     95
     96
     97
     98
     99
    100
    101
    102
    103
    104
    105
    106
    107
    108
    109
    110
    111
    112
    113
    114
    115
    116
    117
    118
    119
    120
    121
    122
    123
    124
    125
    126
    127
    128
    129
    130
    131
    132
    133
    134
    135
    136
    137
    138
    139
    140
    141
    142
    143
    144
    145
    146
    147
    148
    149
    150
    151
    152
    153
    154
    155
    156
    157
    158
    159
    160
    161
    162
    163
    164
    165
    166
    167
    168
    169
    170
    171
    172
    173
    174
    175
    176
    177
    178
    179
    180
    181
    182
    183
    184
    185
    186
    187
    188
    189
    190
    191
    192
    193
    194
    195
    196
    197
    198
    199
    200
    201
    202
    203
    204
    205
    206
    207
    208
    209
    210
    211
    212
    213
    214
    215
    216
    217
    218
    219
    220
    221
    222
    223
    224
    225
    226
    227
    228
    229
    230
    231
    232
    233
    234
    235
    236
    237
    238
    239
    240
    241
    242
    243
    244
    245
    246
    247
    248
    249
    250
    251
    252
    253
    254
    255
    256
    257
    258
    259
    260
    261
    262
    263
    264
    265
    266
    267
    268
    269
    270
    271
    272
    273
    274
    275
    276
    277
    278
    279
    280
    281
    282
    283
    284
    285
    286
    287
    288
    289
    290
    291
    292
    293
    294
    295
    296
    297
    298
    299
    300
    301
    302
    303
    304
    305
    306
    307
    308
    309
    310
    311
    312
    313
    314
    315
    316
    317
    318
    319
    320
    321
    322
    323
    324
    325
    326
    327
    328
    329
    330
    331
    332
    333
    334
    335
    336
    337
    338
    339
    340
    341
    342
    343
    344
    345
    346
    347
    348
    349
    350
    351
    352
    353
    354
    355
    356
    357
    358
    359
    360
    361
    362
    363
    364
    365
    366
    367
    368
    369
    370
    371
    372
    373
    374
    375
    376
    377
    378
    379
    380
    381
    382
    383
    384
    385
    386
    387
    388
    389
    390
    391
    392
    393
    394
    395
    396
    397
    398
    399
    400
    401
    402
    403
    404
    405
    406
    407
    408
    409
    410
    411
    412
    413
    414
    415
    416
    417
    418
    419
    420
    421
    422
    423
    424
    425
    426
    427
    428
    429
    430
    431
    432
    433
    434
    435
    436
    437
    438
    439
    440
    441
    442
    443
    444
    445
    446
    447
    448
    449
    450
    451
    452
    453
    454
    455
    456
    457
    458
    459
    460
    461
    462
    463
    464
    465
    466
    467
    468
    469
    470
    471
    472
    473
    474
    475
    476
    477
    478
    479
    480
    481
    482
    483
    484
    485
    486
    487
    488
    489
    490
    491
    492
    493
    494
    495
    496
    497
    498
    499
    500
    501
    502
    503
    504
    505
    506
    507
    508
    509
    510
    511
    512
    513
    514
    515
    516
    517
    518
    519
    520
    521
    522
    523
    524
    525
    526
    527
    528
    529
    530
    531
    532
    533
    534
    535
    536
    537
    538
    539
    540
    541
    542
    543
    544
    545
    546
    547
    548
    549
    550
    551
    552
    553
    554
    555
    556
    557
    558
    559
    560
    561
    562
    563
    564
    565
    566
    567
    568
    569
    570
    571
    572
    573
    574
    575
    576
    577
    578
    579
    580
    581
    582
    583
    584
    585
    586
    587
    588
    589
    590
    591
    592
    593
    594
    595
    596
    597
    598
    599
    600
    601
    602
    603
    604
    605
    606
    607
    608
    609
    610
    611
    612
    613
    614
    615
    616
    617
    618
    619
    620
    621
    622
    623
    624
    625
    626
    627
    628
    629
    630
    631
    632
    633
    634
    635
    636
    637
    638
    639
    640
    641
    642
    643
    644
    645
    646
    647
    648
    649
    650
    651
    652
    653
    654
    655
    656
    657
    658
    659
    660
    661
    662
    663
    664
    665
    666
    667
    668
    669
    670
    671
    672
    673
    674
    675
    676
    677
    678
    679
    680
    681
    682
    683
    684
    685
    686
    687
    688
    689
    690
    691
    692
    693
    694
    695
    696
    697
    698
    699
    700
   
```

```

Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILKE, BETTINA
APPLICANT: PREAC-MURISIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFENER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

Query Match 34.3%, Score 616.5, DB 4, Length 210;
Best Local Similarity 69.0%, Pred. No. 1.3e-41;
Matches 129, Conservative 24, Mismatches 33, Indels 1, Gaps 1

0Y 1 MACNNSGKDGNAASADSSEYKGPVLTEISKITESTSNNAVVLAKVEETLASIDELATKA 60
Db 17 ISCNNSGDKGNTSANSADSEYKGPVLTEISKITESTSNNAVVLAVLEVEALLSIDELATAKA 76
0Y 61 IGKTI-GNNGLEANOSKNTSLSGAVYASIDLIAEKILNVLKNEELKEKIDTAROCSTFTN 119
Db 77 IGKRIHNNNGLDTEHNHNSLLAGAVYASTLIKOKIDGLKNGLEKIDAKKCSSTFTN 136
0Y 120 KIKSHNAVVLGIDNLTDNNAQRILKKHANKDKGALELELFAVEVLSAAADITLKNAYK 179

```

```

Db 137 K1KEHNTDLGKEGVYTDADAKFAILKTNGTKTKGABELGLPESVEVLSKAEMLANSVK 196
OY 180 ELTSPV 186
| | | | | : |
Db 197 ELTSPV 203

RESULT 9
US-08-158-353-4
; Sequence 4, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UCT93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-158-353-4

Query Match 32.7%; Score 588.5; DB 1; Length 212;
Best Local Similarity 68.8%; Pred. No. 2.1e-39;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3.

OY 1 MACNNSGSDG-NASANSADSEKYNLEISKITEESNAVLAVEVEVTLASIDELATK 59
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 ISCNNGGKGGGSPADNPDESAGKPNLTETISKRIDSNAFVAVEVEVTLASIDELAK 76
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 60 AIGKRI-NNNGSEANOSKNTSLSGAVAIISDILAEKLVLNK-NELTEKIDPAKOCSTEF 117
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 AIGQITDNNNGIALAINONGSILACATYAI STLITERLSKRLNLEELKTIARAKKCSPEF 136
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 118 TNKLKSEHAVIGLMDITDNOARILKLRHANKDGAAELEKLFKAVENLSKAADTLKNA 177
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 TNKLKSGHADGKQDQATDDHAKAAILKTHATYTDGAKEFKDLFESVEGLLKAAYALTVNS 196
OY 178 VKELTSPV 186
| | | | | : |
Db 197 VKELTSPV 205

```

```

      Db      197 VKELSPV 205

      RESULT  11
      US-07-903-580-2
      ; Sequence 2, Application US/07903580
      ; Patent No. 6221363
      " "

```

RESULT 12
US-08-235-836C-34
; Sequence 34, Application US/08235836C
; Patent No. 6246562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.

TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-34

Query Match 32.5%; Score 583.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 5.1e-39;
Matches 129; Conservative 21; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNNGKDG-MASANSADSVKGPNTETSKRTTESNAVAVLVKVEETLSTDELATK 59
DB 17 ISCSNKGKGDSTNPADDSAGPNLTETSKRTTDSNAVLAVKEETVLSTDELAKK 76
60 AIGKKI-GNNGLEANSKNTSLGAYAISDLIAEKLNVLN-BELKEKIDTAKOCSTEF 117
DB 77 AIGOKIDNNNGLAALNNQNSLAGAYAI STLITERKLSKLNLEELKTELAKKKEEF 136
QY 118 TNKLRSEHAVLGIDNLTDDNAQRALIKKHANKDKGAEELEKLEKRAVENLSKAADTLKNA 177
DB 137 TNKLRSGHADLGKODATDDHAKAAIKLTHATTDGKAKEFKDLFESVGLLKAQVALTNS 196
QY 178 VKELTSPV 186
DB 197 VKELTSPV 205

RESULT 13
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Sautscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
TITLE OF INVENTION: burgdorferi
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293

CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match 32.2%; Score 578.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 1.3e-38;
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

QY 1 MACNNGKDG-MASANSADSVKGPNTETSKRTTESNAVAVLVKVEETLSTDELATK 59
DB 17 ISCSNKGKGLTSTNPADDSAGPNLTETSKRTTDSNAVLAVKEETVLSTDELAKK 76
QY 60 AIGKKI-GNNGLEANSKNTSLGAYAISDLIAEKLNVLN-BELKEKIDTAKOCSTEF 117
DB 77 AIGOKIDNNNGLAALNNQNSLAGAYAI STLITERKLSKLNLEELKTELAKKKEEF 136
QY 118 TNKLRSEHAVLGIDNLTDDNAQRALIKKHANKDKGAEELEKLEKRAVENLSKAADTLKNA 177
DB 137 TNKLRSGHADLGKODATDDHAKAAIKLTHATTDGKAKEFKDLFESVGLLKAQVALTNS 196
QY 178 VKELTSPV 186
DB 197 VKELTSPV 205

RESULT 14
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSEC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDOFFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990

APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 212
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDOFFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE:
US-08-209-603E-11

Query Match 32.2%; Score 578.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 1.3e-38;
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;
QY 1 MACNNGSKDGS-NASANSADSVKGPMLTEISKITESNAVLAKEVEETLLASIDELATK 59
DB 17 ISCNNGSGKGIITSTNPADSAKGPMLTEISKITDSNAFLAVKEVETLVISIDELAKK 76
QY 60 AIGKRT-GNNGEANSKNTSLSGAYAISDLIAEKLNVLNK-EELKEKIDPAKOCSTEF 117
DB 77 AIGKRTDNNNGALANNONGSLAGYAIISTLTETKSLKLNLEELKEITIAAKKCSSEF 136
QY 118 TNKLSSEHAVLGDINTDNOARAILKKHANKDGALELEKLFKAVENTSKAAODTLKNA 177
DB 137 TNKLSGSHADLGKODATDDHAKAAILKTHATTDGAKKEKDLFESVEGLTKAAQVALTNS 196
178 VKELTSPIV 186
197 VKELTSPV 205

RESULT 15
US-08-235-836C-36
Sequence 36, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
NUMBER OF INVENTION: Borrelia Polypeptides and Uses Therefor
CORRESPONDENCE ADDRESS: 144
ADDRESS: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BMD93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-36

Query Match 31.3%; Score 563; DB 4; Length 207;
Best Local Similarity 62.6%; Pred. No. 2e-37;
Matches 117; Conservative 34; Mismatches 32; Indels 4; Gaps 3;
QY 1 MACNNGSKDGS-NASANSADSVKGPMLTEISKITESNAVLAKEVEETLLASIDELATK 60
DB 17 ISCNNGSG--GDSASTNPESAKGPMLTVISKITDSNAFLAVKEVEETLLASIDEL--SKA 73
QY 61 IGKRTGNG-LEANSKNTSLSGAYAISDLIAEKLNVLNK-EELKEKIDPAKOCSTEF 119
DB 74 IGKRTKNDGTDNEANRNESLIAGAYEISKLITQKLSVNSEELKKIKKEAKKOCSEFTT 133
QY 120 KLSSEHAVLGDINTDNOARAILKKHANKDGALELEKLFKAVENTSKAAODTLKNA 179
DB 134 KLSSEHAVLGDINTDNOARAILKKHANKDGALELEKLFKAVENTSKAAODTLKNA 193
QY 180 ELTSPIV 186
DB 194 ELTNPV 200

Search completed: March 18, 2002, 09:55:35
Job time: 348 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:45 ; Search time 621.2 Seconds
(without alignments)
164.485 Million cell updates/sec

Title: US-09-596-746A-38

Perfect score: 1797

Sequence: 1 MACNNSGKDDGNASANSADSE.....ENLSKAQDTLKNAVKETLS 368

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending-Patents-AA-Main:*
1: /cgn2_6/ptodata/2/paa/PCUTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1797	100.0	368	19	US-09-596-746-38
2	1797	100.0	368	19	US-09-596-746A-38
3	1789	99.6	391	19	US-09-596-746-76
4	1789	99.6	392	19	US-09-596-746A-76
5	1538	85.6	368	19	US-09-596-746A-42
6	1533	85.3	367	19	US-09-596-746-42
7	1530	85.1	391	19	US-09-596-746-74
8	1530	85.1	392	19	US-09-596-746A-74
9	1490.5	82.9	369	19	US-09-596-746A-40

10	1485.5	82.7	368	19	US-09-596-746-40	Sequence 40, Appl
11	1484	82.6	370	19	US-09-596-746A-32	Sequence 32, Appl
12	1482.5	82.5	392	19	US-09-596-746-72	Sequence 72, Appl
13	1482.5	82.5	393	19	US-09-596-746A-72	Sequence 72, Appl
14	1479	82.3	369	19	US-09-596-746-32	Sequence 32, Appl
15	1476	82.1	393	19	US-09-596-746-68	Sequence 68, Appl
16	1476	82.1	394	19	US-09-596-746A-68	Sequence 68, Appl
17	1429	68.4	374	19	US-09-596-746A-26	Sequence 26, Appl
18	1429	68.4	377	19	US-09-596-746A-30	Sequence 30, Appl
19	1225.5	68.2	377	19	US-09-596-746A-30	Sequence 30, Appl
20	1225	68.2	370	19	US-09-596-746A-36	Sequence 36, Appl
21	1224	68.1	373	19	US-09-596-746-26	Sequence 26, Appl
22	1221	67.9	393	19	US-09-596-746-62	Sequence 62, Appl
23	1221	67.9	398	19	US-09-596-746A-62	Sequence 62, Appl
24	1220	67.9	369	19	US-09-596-746-36	Sequence 36, Appl
25	1219.5	67.9	400	19	US-09-596-746-82	Sequence 82, Appl
26	1219.5	67.9	401	19	US-09-596-746A-82	Sequence 82, Appl
27	1217	67.7	393	19	US-09-596-746-66	Sequence 66, Appl
28	1217	67.7	394	19	US-09-596-746A-66	Sequence 66, Appl
29	1181.5	65.7	375	19	US-09-596-746A-24	Sequence 24, Appl
30	1178	65.6	378	19	US-09-596-746A-28	Sequence 28, Appl
31	1178	65.6	378	19	US-09-596-746A-28	Sequence 28, Appl
32	1177.5	65.5	371	19	US-09-596-746A-34	Sequence 34, Appl
33	1176.5	65.5	374	19	US-09-596-746-24	Sequence 24, Appl
34	1173.5	65.3	398	19	US-09-596-746-60	Sequence 60, Appl
35	1173.5	65.3	399	19	US-09-596-746A-60	Sequence 60, Appl
36	1172.5	65.2	370	19	US-09-596-746-34	Sequence 34, Appl
37	1172	65.2	401	19	US-09-596-746-80	Sequence 80, Appl
38	1172	65.2	402	19	US-09-596-746A-80	Sequence 80, Appl
39	1169.5	65.1	394	19	US-09-596-746-64	Sequence 64, Appl
40	1169.5	65.1	395	19	US-09-596-746A-64	Sequence 64, Appl
41	1167.5	65.0	560	19	US-09-596-746A-52	Sequence 52, Appl
42	1167.5	65.0	560	19	US-09-596-746A-52	Sequence 52, Appl
43	1162.5	64.7	385	19	US-09-596-746A-56	Sequence 56, Appl
44	1157.5	64.4	384	19	US-09-596-746-56	Sequence 56, Appl
45	1154.5	64.2	408	19	US-09-596-746-78	Sequence 78, Appl

ALIGNMENTS

RESULT 1
US-09-596-746-38
Sequence 38, Application US/09596746
GENERAL INFORMATION:
APPLICANT: Datwyler, Raymond J.
APPLICANT: Seinst, Gerald
APPLICANT: Dykhizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease In Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38
LENGTH: 368
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746-38

Query Match 100.0%; Score 1797; DB 19; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MACNNSGKDDGNASANSADSEVKGNPLTEISKKTIESNAVLAWEVETLLASIDELATKA 60


```
RESULT 4
US-09-596-746A-76
; Sequence 76, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
EO ID NO 76
LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-76

Query Match          99.6%; Score 1789; DB 19; Length 392;
Best Local Similarity 99.7%; Pred. No. 6.1e-124;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGNASANSADSVKGPNTLEISKITESNAVYLAVKEVETLLASIDELATKA 61
DB 26 SCNNSGKDGNASANSADSVKGPNTLEISKITESNAVYLAVKEVETLLASIDELATKA 85
QY 62 GKRIKNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOCSTEFNTK 121
DB 86 GKRIKNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOCSTEFNTK 145
QY 122 KSHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDTLKNAYKE 181
DB 146 KSHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDTLKNAYKE 205
QY 182 TSPVHGNNSRKDGNASTNSADESVKGPNTLEISKITESNAVYLAVKEVETLLASIDEL 241
DB 206 TSPVHGNNSRKDGNASTNSADESVKGPNTLEISKITESNAVYLAVKEVETLLASIDEL 265
QY 242 ATKAIGKIKGNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOCSTE 301
DB 266 ATKAIGKIKGNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOCSTE 325
QY 302 FTKKLSEHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDTLKN 361
DB 326 FTKKLSEHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDTLKN 385
QY 362 AVKEKLS 368
DB 386 AVKEKLS 392

RESULT 5
US-09-596-746A-42
; Sequence 42, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
```

```
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ospc Chimera
US-09-596-746A-42

Query Match          85.6%; Score 1538; DB 19; Length 368;
Best Local Similarity 86.7%; Pred. No. 2.1e-105;
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

QY 1 MACNNSGKDGNASANSADSVKGPNTLEISKITESNAVYLAVKEVETLLASIDELATKA 60
DB 1 MACNNSGKDGNASANSADSVKGPNTLEISKITESNAVYLAVKEVETLLASIDELATKA 60
QY 61 GKRIKNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOCSTEFNTK 120
DB 61 GKRIKNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOCSTEFNTK 120
QY 121 KSHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDTLKNAYKE 180
DB 121 KSHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDTLKNAYKE 180
QY 181 TSPVHGNNSRKDGNASTNSADESVKGPNTLEISKITESNAVYLAVKEVETLLASIDE 240
DB 181 TSPVHGNNSRKDGNASTNSADESVKGPNTLEISKITESNAVYLAVKEVETLLASIDE 240
QY 241 LATKAIGKIKGNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOC 298
DB 241 LATKAIGKIKGNGGLEANOSKNTSLSGAYAISDLIAEKLNVKNELEKIDTAQOC 300
QY 299 STEFTNKLKSEHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDT 358
DB 301 SEDFTNKLKSEHAVLGLDNLTDNAORAILKHKHANKDGALEKLFRAVENLSKAADDT 360
QY 359 LKNAYKE 366
DB 361 LANSYKE 368

RESULT 6
US-09-596-746-42
; Sequence 42, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinozt, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ospc Chimera
US-09-596-746-42

Query Match          85.3%; Score 1533; DB 19; Length 367;
```

Best Local Similarity 86.6%; Pred. No. 4.9e-105;
Matches 317; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

```
OY 2 ACNNSGKDGASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
Db 1 ACNNSGKDGASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 60
OY 62 GKIGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNKL 121
Db 61 GKIGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNKL 120
OY 122 KSEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTLKNAVKEL 181
Db 121 KSEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTLKNAVKEL 180
OY 182 TSPYHGNNSRKDGASNTSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
Db 181 TSPYHGNNSRKDGASNTSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 240
OY 242 ATKAIGKKI-GNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOC 299
Db 241 ATKAIGKKIQOONGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSSEKLEKENNAKKS 300
OY 300 TETTNKLESEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTL 359
Db 301 EPTFKLEGEHQGLIENVTDENAKKAILITDAADKGAEELEKFKAVENLSKAQAKEM 360
OY 360 KNAVKEL 366
Db 361 ANSVKEL 367
```

RESULT 7
US-09-596-746-74

```
; Sequence 74, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 391
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-74
```

Query Match 85.1%; Score 1530; DB 19; Length 391;
Best Local Similarity 86.4%; Pred. No. 8.9e-105;
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

```
OY 2 ACNNSGKDGASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
Db 25 SCNNSGKDGASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 84
OY 62 GKIGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNKL 121
Db 85 GKIGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNKL 144
OY 122 KSEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTLKNAVKEL 181
Db 145 KSEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTLKNAVKEL 204
```

```
OY 182 TSPYHGNNSRKDGASNTSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
Db 205 TSPYHGNNSRKDGASNTSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 264
OY 242 ATKAIGKKI-GNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOC 299
Db 265 ATKAIGKKIQOONGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSSEKLEKENNAKKS 324
OY 300 TETTNKLESEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTL 359
Db 325 EPTFKLEGEHQGLIENVTDENAKKAILITDAADKGAEELEKFKAVENLSKAQAKEM 384
OY 360 KNAVKEL 366
Db 385 ANSVKEL 391
```

RESULT 8
US-09-596-746A-74

```
; Sequence 74, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596/746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-74
```

Query Match 85.1%; Score 1530; DB 19; Length 392;
Best Local Similarity 86.4%; Pred. No. 8.9e-105;
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

```
OY 2 ACNNSGKDGASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
Db 26 SCNNSGKDGASANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 85
OY 62 GKIGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNKL 121
Db 86 GKIGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOCSTFTNKL 145
OY 122 KSEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTLKNAVKEL 181
Db 146 KSEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTLKNAVKEL 205
OY 182 TSPYHGNNSRKDGASNTSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
Db 206 TSPYHGNNSRKDGASNTSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 265
OY 242 ATKAIGKKI-GNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNEELKEKIDTAQOC 299
Db 266 ATKAIGKKIQOONGGLAVEAGHNGTLLAGAYTTSKLTOKLDGKNSSEKLEKENNAKKS 325
OY 300 TETTNKLESEHAVLGLDNLTDNAQRALIKKHANKDGAEELEKFKAVENLSKAQDTL 359
Db 326 EPTFKLEGEHQGLIENVTDENAKKAILITDAADKGAEELEKFKAVENLSKAQAKEM 385
OY 360 KNAVKEL 366
Db 386 ANSVKEL 392
```


US-09-596-746a-32

Query Match 82.6%; Score 1484; DB 19; Length 370;
 Best Local Similarity 83.8%; Pred. No. 2.1e-101;
 Matches 311; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

QY 1 MACNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 60
 DB 1 MACNSGKDGNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 59
 QY 61 IGKRTGNNG-LEANSOKMTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOCSTFT 118
 DB 60 IGKRTKNGSLDNEANRNEESLAGAVTITSLITQKLKSLNGSEGLKEKIAAKKCESEFS 119
 QY 119 NKLSEHAVIGLIDNLTDDNAORAILKKH-ANKDKGALELEKLFRAVENLSKAODTLKNA 177
 DB 120 TKLKNHQAQGLQGVTDENAKKAILKNAAGKDKGVELEKLSLSSLSKAAKEMLANS 179
 QY 178 VKELTSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLAS 237
 DB 180 VKELTSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLAS 239
 QY 238 IDELATRAIGKRTGNNGLEANSOKMTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKO 297
 DB 240 IDELATRAIGKRTGNNGLEANSOKMTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKO 299
 QY 298 CSTFTNKLKSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAOD 357
 DB 300 CSTFTNKLKSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAOD 359
 QY 358 TLKNAVKELTS 368
 DB 360 TLKNAVKELTS 370

RESULT 12
 US-09-596-746-72
 ; Sequence 72, Application US/09596746
 ; GENERAL INFORMATION:
 ; APPLICANT: Datwyler, Raymond J.
 ; APPLICANT: Seinoz, Gerald
 ; APPLICANT: Dykhuzen, Daniel
 ; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; TITLE OF INVENTION: Borrelia atzei That Cause Lyme Disease in Humans
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746
 ; PRIOR FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 72
 ; LENGTH: 392
 ; TYPE: PRT
 ; ORGANISM: ospc Chimera
 US-09-596-746-72

Query Match 82.5%; Score 1482.5; DB 19; Length 392;
 Best Local Similarity 84.0%; Pred. No. 2.9e-101;
 Matches 310; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNSGKGNNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 61
 DB 25 SCNSGKGNNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 84
 QY 62 GKRTGNNGLEANSOKMTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTFTNKL 121
 DB 85 GKRTGNNGLEANSOKMTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTFTNKL 144

QY 122 KSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTLKNAVKEL 181
 DB 145 KSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTLKNAVKEL 204
 QY 182 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 241
 DB 205 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 264
 QY 242 ATTAIGKRTGNNG-LEANSOKMTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOC 299
 DB 265 A-RAIGKRTGNNG-LEANSOKMTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOC 323
 QY 300 TEFNTNKLKSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTL 359
 DB 324 TEFNTNKLKSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTL 383
 QY 360 KNAVKELTS 368
 DB 384 TNSVKELTS 392

RESULT 13
 US-09-596-746a-72
 ; Sequence 72, Application US/09596746a
 ; GENERAL INFORMATION:
 ; APPLICANT: Datwyler, Raymond J.
 ; APPLICANT: Seinoz, Gerald
 ; APPLICANT: Dykhuzen, Daniel
 ; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; TITLE OF INVENTION: Borrelia atzei That Cause Lyme Disease in Humans
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746a
 ; PRIOR FILING DATE: 2000-06-19
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 72
 ; LENGTH: 393
 ; TYPE: PRT
 ; ORGANISM: ospc Chimera
 US-09-596-746a-72

Query Match 82.5%; Score 1482.5; DB 19; Length 393;
 Best Local Similarity 84.0%; Pred. No. 2.9e-101;
 Matches 310; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

QY 2 ACNSGKGNNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 61
 DB 26 SCNSGKGNNSANSADSVKGPMLTEISKRTTESNAVLAKEVEETLLASIDELATKA 85
 QY 62 GKRTGNNGLEANSOKMTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTFTNKL 121
 DB 86 GKRTGNNGLEANSOKMTSLSGAVAIISDLIAEKLNVLKNEELKEKIDTAKOCSTFTNKL 145
 QY 122 KSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTLKNAVKEL 181
 DB 146 KSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTLKNAVKEL 205
 QY 182 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 241
 DB 206 TSPYVHGNNSRKDGNASTNSADESVKGPMLTEISKRTTESNAVLAKEVEETLLASIDEL 265
 QY 242 ATTAIGKRTGNNG-LEANSOKMTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOC 299
 DB 266 A-RAIGKRTGNNG-LEANSOKMTSLSGAVAIISDLIAEKLNVLKNE-LEKEIDTAKOC 324
 QY 300 TEFNTNKLKSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTL 359
 DB 325 TEFNTNKLKSEHAVIGLIDNLTDDNAORAILKKHANKDKGALELEKLFRAVENLSKAODTL 384

QY 360 KNAVKELTS 368
DB 385 TNSVKELTS 393

RESULT 14

US-09-596-746-32
Sequence 32, Application US/09596746
GENERAL INFORMATION:
APPLICANT: Dettwyler, Raymond J.
APPLICANT: Selnost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 369
TYPE: PRN
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ospc Chimera
US-09-596-746-32

Query Match 82.3%; Score 1479; DB 19; Length 369;
Best Local Similarity 83.8%; Pred. No. 4.9e-101;
Matches 310; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDNASANSADSVKGNLTETISKTTESNAVLAKEVETLLASIDELATKAI 61
DB 1 ACNNSGKDNSTANSADSVKGNLTETISKTTESNAVLAKEVETLLASIDELATKAI 59
QY 62 GKRIKNGG-LEANOSKNTSLSGAVAIISDLIAEKLNVKNEE-LKEKIDTAKOCSTEFN 119
DB 60 GKRIKNGGSLDNANRNEISLAGAYITSTLTOKLSKLNCEGLKEKIAAKKCESEFST 119
QY 120 KLSSEHAVGLDNLTDNAORAILKKH-ANKDGAELKFRVAVENLSKAADDTLKNV 178
DB 120 KLDNHAQGLIGVTDENAKKAILKANAAGKDGVELEKLSLSLAKAEMLANSV 179
QY 179 KELTSPIVHGNNRSKDNASTNSADSVKGNLTETISKTTESNAVLAKEVETLLASI 238
DB 180 KELTSPIVHGNNRSKDNASTNSADSVKGNLTETISKTTESNAVLAKEVETLLASI 239
QY 239 DELATKAIGKRIKNGGLEANOSKNTSLSGAVAIISDLIAEKLNVKNEE-LKEKIDTAKOC 298
DB 240 DELATKAIGKRIKNGGLEANOSKNTSLSGAVAIISDLIAEKLNVKNEE-LKEKIDTAKOC 299
QY 299 STEFTNKLKSEHAVGLDNLTDNAORAILKKHANKDGAELKFRVAVENLSKAADDT 358
DB 300 STEFTNKLKSEHAVGLDNLTDNAORAILKKHANKDGAELKFRVAVENLSKAADDT 359
QY 359 LKNAVKELTS 368
DB 360 LKNAVKELTS 369

RESULT 15

US-09-596-746-68
Sequence 68, Application US/09596746
GENERAL INFORMATION:
APPLICANT: Dettwyler, Raymond J.
APPLICANT: Selnost, Gerald
APPLICANT: Dykhuizen, Daniel

APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 68
LENGTH: 393
TYPE: PRN
ORGANISM: ospc Chimera
US-09-596-746-68

Query Match 82.1%; Score 1476; DB 19; Length 393;
Best Local Similarity 83.5%; Pred. No. 8.9e-101;
Matches 309; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDNASANSADSVKGNLTETISKTTESNAVLAKEVETLLASIDELATKAI 61
DB 25 ACNNSGKDNSTANSADSVKGNLTETISKTTESNAVLAKEVETLLASIDELATKAI 83
QY 62 GKRIKNGG-LEANOSKNTSLSGAVAIISDLIAEKLNVKNEE-LKEKIDTAKOCSTEFN 119
DB 84 GKRIKNGGSLDNANRNEISLAGAYITSTLTOKLSKLNCEGLKEKIAAKKCESEFST 143
QY 120 KLSSEHAVGLDNLTDNAORAILKKH-ANKDGAELKFRVAVENLSKAADDTLKNV 178
DB 144 KLDNHAQGLIGVTDENAKKAILKANAAGKDGVELEKLSLSLAKAEMLANSV 203
QY 179 KELTSPIVHGNNRSKDNASTNSADSVKGNLTETISKTTESNAVLAKEVETLLASI 238
DB 204 KELTSPIVHGNNRSKDNASTNSADSVKGNLTETISKTTESNAVLAKEVETLLASI 263
QY 239 DELATKAIGKRIKNGGLEANOSKNTSLSGAVAIISDLIAEKLNVKNEE-LKEKIDTAKOC 298
DB 264 DELATKAIGKRIKNGGLEANOSKNTSLSGAVAIISDLIAEKLNVKNEE-LKEKIDTAKOC 323
QY 299 STEFTNKLKSEHAVGLDNLTDNAORAILKKHANKDGAELKFRVAVENLSKAADDT 358
DB 324 STEFTNKLKSEHAVGLDNLTDNAORAILKKHANKDGAELKFRVAVENLSKAADDT 383
QY 359 LKNAVKELTS 368
DB 384 LKNAVKELTS 393

Search completed: March 18, 2002, 10:08:45
Job time: 973 sec

Tue Mar 19 10:57:53 2002

us-09-596-746a-38.rapm

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw.model

Run on: March 18, 2002, 09:58:23 ; Search time 81.02 seconds

(Without alignments)
290.158 Million cell updates/sec

Title: US-09-596-746a-38

Perfect score: 1797
Sequence: 1 MACNNSGCDGNASANSADSE.....ENLSKRAODTLKNAYKELTS 368

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Pending Patents AA New: *
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep: *
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	588.5	32.7	212	6	US-09-974-992-7 Sequence 7, Appl1
2	583	32.4	211	6	US-09-974-992-5 Sequence 5, Appl1
3	567	31.6	207	6	US-09-974-992-3 Sequence 3, Appl1
4	561	30.6	837	6	US-09-815-242-5883 Sequence 5883, Ap
5	561	30.6	837	7	US-10-072-851-5883 Sequence 5883, Ap
6	561	30.6	837	6	US-09-815-242-13080 Sequence 13080, A
7	561	30.6	837	7	US-10-072-851-13080 Sequence 13080, A
8	561	30.6	837	6	US-09-815-242-5835 Sequence 5835, Ap
9	561	30.6	837	6	US-10-072-851-5835 Sequence 5835, Ap
10	561	30.6	837	6	US-09-815-242-12996 Sequence 12996, A
11	561	30.6	837	7	US-10-072-851-12996 Sequence 12996, A
12	561	30.6	837	7	US-10-032-585-7646 Sequence 7646, Ap
13	561	30.6	837	7	US-10-072-851-15590 Sequence 15590, A
14	561	30.6	837	6	US-09-815-242-5815 Sequence 5815, Ap
15	561	30.6	837	6	US-10-072-851-5815 Sequence 5815, Ap
16	561	30.6	837	6	US-09-815-242-12913 Sequence 12913, A
17	561	30.6	837	6	US-10-072-851-12913 Sequence 12913, A
18	561	30.6	837	6	US-09-815-242-15045 Sequence 15045, A
19	561	30.6	837	6	US-09-815-242-15045 Sequence 15045, A
20	561	30.6	837	6	US-09-815-242-15044 Sequence 15044, A
21	561	30.6	837	6	US-09-815-242-13229 Sequence 13229, A
22	561	30.6	837	6	US-09-815-242-13228 Sequence 13228, A
23	561	30.6	837	6	US-09-815-242-13227 Sequence 13227, A
24	561	30.6	837	6	US-09-815-242-13227 Sequence 13227, A
25	561	30.6	837	6	US-09-815-242-13227 Sequence 13227, A

26	162.5	9.0	1269	6	US-09-708-427-19881	Sequence 19881, A
27	162.5	9.0	2056	6	US-09-614-150-4824	Sequence 4824, Ap
28	159	8.8	2437	6	US-09-815-242-5834	Sequence 5834, Ap
29	159	8.8	2437	7	US-10-072-851-5834	Sequence 5834, Ap
30	158	8.8	5795	6	US-09-815-242-12610	Sequence 12610, A
31	158	8.8	5795	7	US-10-072-851-12610	Sequence 12610, A
32	157.5	8.8	2086	6	US-09-815-242-5639	Sequence 5639, Ap
33	157.5	8.8	2086	7	US-10-072-851-5639	Sequence 5639, Ap
34	157	8.7	1639	6	US-09-614-150-8213	Sequence 8213, Ap
35	156.5	8.7	1013	6	US-09-614-150-9139	Sequence 9139, Ap
36	152.5	8.5	1128	7	US-10-032-585-7627	Sequence 7627, Ap
37	152.5	8.5	1128	7	US-10-072-851-15571	Sequence 15571, A
38	152	8.5	2025	6	US-09-815-242-5703	Sequence 5703, Ap
39	152	8.5	2025	7	US-10-072-851-5703	Sequence 5703, Ap
40	152	8.5	3158	6	US-09-815-242-12611	Sequence 12611, A
41	152	8.5	3158	7	US-10-072-851-12611	Sequence 12611, A
42	151.5	8.4	2013	6	US-09-614-150-13758	Sequence 13758, A
43	149.5	8.3	1765	7	US-10-037-182-8	Sequence 8, Appl1
44	149.5	8.3	1786	5	US-09-938-275-6	Sequence 6, Appl1
45	149.5	8.3	1786	6	US-09-873-676-113	Sequence 113, Appl

ALIGNMENTS

RESULT 1
US-09-974-992-7
Sequence 7, Application US/09974/992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Thiesen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OSpC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09/974, 992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180, 089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia afzelii
US-09-974-992-7

Query Match 32.7%, Score 588.5; DB 6; Length 212;
Best Local Similarity 68.8%; Pred. No. 2.9e-32;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNNSGCDG-NASANSADSEYKGNLTREISKRTESNAVYIAVEVETLTLASIDELATK 59
DB 17 ISCNNSGCGDSASTNPDESAKGNLTREISKRTIDSNFAVIAVEVETLTLASIDELATK 76
QY 60 AIGKRI-GNNGLEAENSQNTSLSGAVYISDIAELKLVNKN-EELKEKIDTAKCSTEE 117
DB 77 AIGKIDNNNGIALNNNGSLAGAVYISTITREKLRKLEDEKTEIAKAKKCEEF 136
QY 118 TNKLSKSHADIGKQADTDHAKAALIKTHATGDKAKFEKDLFESEVGLTKAQAVALTNS 196
DB 137 TNKLSKSHADIGKQADTDHAKAALIKTHATGDKAKFEKDLFESEVGLTKAQAVALTNS 196
QY 178 VKELTSPV 186
DB 197 VKELTSPV 205
RESULT 2
US-09-974-992-5

Query Match	31.6%	Score 567;	DB 6;	Length 207;
Best Local Similarity	63.1%	Pred. No. 7.6e-31;		
Matches 118; Conservative	33;	Mismatches 32;	Indels 4;	Gaps 3;

[illegible]

OY 165 NLS--KAAODTLKNAVKELTSPIVHGNNSRKDNASTNADSEYVGPULTEISKITESEN 222
DB 273 NADALNTAMTILKNGIOD-----QNTIKOG--VNFTDADBA-----KRN 309
OY 223 AVLAIVEVEFTLL-----ASIDELATKAIGKIGNNGLEANOSKNTSLSGAVAIISDL 275
DB 310 AVTNAVTOAEOLINKAOGPNTSKDGVETALENVORAKNELNGONVANAKTTAKALNML 369
OY 276 IAEKLVNLKNELEKID-----TAKOCSTEFNKLKSENAVLGLDN-LTDD 321
DB 370 TS--INNOKKALKSQIEGATTVAGVNOVSTTASELNTAMSN-----LONGINDE 417
OY 322 NAORAILKHKNDKGALEKLFKAVENLSKAAODTLKNAVKELTS 368
DB 418 AATKALNGTONLEKAKOHANTAIIDGLSHLTNAOKKALKOLYOOSTT 464

RESULT 5

Sequence 5883, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyekind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5883
LENGTH: 837
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-072-851-5883

Query Match 10.6%; Score 191; DB 7; Length 837;
Best Local Similarity 22.4%; Pred. No. 3.8e-05;
Matches 91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;
OY 4 NNSGDKGNASNSADE---SVKGPULTEISKITESEMANV--VLAVKE-VELTLASI----- 53
DB 100 NNNLEVAKONANTAIIDGLTSLNGPQAKLKEOVGATTLPNQVTRDNAQTNTMTMKGLR 159
OY 54 DELATKAIGKIGNNGLEANOSKNTSLSGAVAIISDLIAE---KLNVLKNELEKIDT 109
DB 160 DSIANEAT-IRAGQNYTDAONKOTDYSAYTAAKAIIGOTTSPEMNOEINOAKDY-T 217
OY 110 AKOCSTFTNKLK-----SENAVLGLDNLTDDNAORAILKHKNDKGALEKLFKAVE 164
DB 218 AKOQALNGOENLRTAQTNKQHLNGLSLDITD--AQKDAVKROI---EGATHVNEVTOAON 272
OY 165 NLS--KAAODTLKNAVKELTSPIVHGNNSRKDNASTNADSEYVGPULTEISKITESEN 222
DB 273 NADALNTAMTILKNGIOD-----QNTIKOG--VNFTDADBA-----KRN 309
OY 223 AVLAIVEVEFTLL-----ASIDELATKAIGKIGNNGLEANOSKNTSLSGAVAIISDL 275

DB 310 AVTNAVTOAEOLINKAOGPNTSKDGVETALENVORAKNELNGONVANAKTTAKALNML 369
OY 276 IAEKLVNLKNELEKID-----TAKOCSTEFNKLKSENAVLGLDN-LTDD 321
DB 370 TS--INNOKKALKSQIEGATTVAGVNOVSTTASELNTAMSN-----LONGINDE 417
OY 322 NAORAILKHKNDKGALEKLFKAVENLSKAAODTLKNAVKELTS 368
DB 418 AATKALNGTONLEKAKOHANTAIIDGLSHLTNAOKKALKOLYOOSTT 464

RESULT 6

US-09-815-242-13080
Sequence 13080, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyekind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13080
LENGTH: 875
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-13080

Query Match 10.6%; Score 191; DB 6; Length 875;
Best Local Similarity 22.4%; Pred. No. 4e-05;
Matches 91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;
OY 4 NNSGDKGNASNSADE---SVKGPULTEISKITESEMANV--VLAVKE-VELTLASI----- 53
DB 137 NNNLEVAKONANTAIIDGLTSLNGPQAKLKEOVGATTLPNQVTRDNAQTNTMTMKGLR 196
OY 54 DELATKAIGKIGNNGLEANOSKNTSLSGAVAIISDLIAE---KLNVLKNELEKIDT 109
DB 197 DSIANEAT-IRAGQNYTDAONKOTDYSAYTAAKAIIGOTTSPEMNOEINOAKDY-T 254
OY 110 AKOCSTFTNKLK-----SENAVLGLDNLTDDNAORAILKHKNDKGALEKLFKAVE 164
DB 255 AKOQALNGOENLRTAQTNKQHLNGLSLDITD--AQKDAVKROI---EGATHVNEVTOAON 309
OY 165 NLS--KAAODTLKNAVKELTSPIVHGNNSRKDNASTNADSEYVGPULTEISKITESEN 222
DB 310 NADALNTAMTILKNGIOD-----QNTIKOG--VNFTDADBA-----KRN 346
OY 223 AVLAIVEVEFTLL-----ASIDELATKAIGKIGNNGLEANOSKNTSLSGAVAIISDL 275

DB 347 AYTNAVTOAQEOLINKAQGPNTSKDGVETALENVQRAKNEELNGNONVANAKTTAKNALNL 406
QY 276 IAEKLVNKEELKEID-----TAKOCSTEFNKLSSEHAVLGLD-LTDD 321
DB 407 TS--INNAQKALKSQIEGATTVAGVNOVSTFASSELNTAMS-----LONGINDE 454
QY 322 NQORALTKHANKDGALEKLEFKAVENLSKAODTLKNAVLELTS 368
DB 455 AATKALNLTQNLKAKQHANTAIIDGLSHLTNAQKELKQVQOSTT 501

RESULT 7

US-10-072-851-13080
Sequence 13080, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13080
LENGTH: 875
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-072-851-13080

Query Match 10.6%; Score 191; DB 7; Length 875;
Best Local Similarity 22.4%; Pred. No. 4e-05;
Matches 91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;

DB 4 NNSGKDNANSANSADE---SVKGNLTETSKITTESNAV--VLAKE-VETLLASI----- 53
DB 137 NNNLEVAKQNNANTAIIDGLTSLNGQKAKLEQVGQATTLTNNQVVRNNAOTLTNTAKMGLR 156
QY 54 DELATKAIGKKNNGLEANSKNTSLSGAYALSIDLIAE---KLVNKEELKEIDT 109
DB 197 DSJANEAT-IKAGQNTYDASQNKQTDYNSAVTAKAALIGQTSPSMAQEIINQAKQV-T 254
QY 110 AKOCSTEFNTKLK-----SEHAVLGLDNLTDNAQRAILKHKHANKDGALEKLEFAYE 164
DB 255 AKQOALNGQENLFTAQTNKQHLNGLSLDLD--AQKDAVKRQI--EGATHVNEVTOAQN 309
QY 165 NLS--KAADOTLKNAVEKLTSPYHGNNSRKDNASTNSADESVKGNLTETSKITTESN 222
DB 310 NADLNTAMTNTKNGID-----QNTIKQS-VNFTDDEA-----KRN 346
QY 223 AVVLAKEVEETLL-----ASIDELATKAIGKKNNGLEANSKNTSLSGAYALSIDL 275
DB 347 AYTNAVTOAQEOLINKAQGPNTSKDGVETALENVQRAKNEELNGNONVANAKTTAKNALNL 406
QY 276 IAEKLVNKEELKEID-----TAKOCSTEFNKLSSEHAVLGLD-LTDD 321
DB 407 TS--INNAQKALKSQIEGATTVAGVNOVSTFASSELNTAMS-----LONGINDE 454

QY 322 NQORALTKHANKDGALEKLEFKAVENLSKAODTLKNAVLELTS 368
DB 455 AATKALNLTQNLKAKQHANTAIIDGLSHLTNAQKELKQVQOSTT 501

RESULT 8

US-09-815-242-5835
Sequence 5835, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5835
LENGTH: 2434
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5835

Query Match 10.1%; Score 182; DB 6; Length 2434;
Best Local Similarity 22.2%; Pred. No. 0.00057;
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;

QY 4 NNSGKDNANSANSADE---SVKGNLTETSKITTESNAV--VLAKE-VETLLASI----- 53
DB 702 NNNLEVAKQNNANTAIIDGLTSLNGQKAKLEQVGQATTLTNNQVVRNNAOTLTNTAKMGLR 761
QY 54 DELATKAIGKKNNGLEANSKNTSLSGAYALSIDLIAE---KLVNKEELKEIDT 109
DB 762 DSJANEAT-IKAGQNTYDASQNKQTDYNSAVTAKAALIGQTSPSMAQEIINQAKQV-T 819
QY 110 AKOCSTEFNTKLK-----SEHAVLGLDNLTDNAQRAILKHKHANKDGALEKLEFAYE 157
DB 820 AKQOALNGQENLFTAQTNKQHLNGLSLDLD--AQKDAVKRQIEGATHVNEVTOAQNND 877
QY 158 KLFKAVENLSKAAD--TLK-----NAVKELTSPY--VHGNNSKRD-- 194
DB 878 ALNTAMTNTKNGIDQNTIKQGVNFTDDEAKRHAATYNTAVTOAQEOLINKAQGPNTSKDGV 937
QY 195 -----GNASTNSADESVKGN--LTETSKITTES-----NA 223
DB 938 ETALENVQRAKNEELNGNONVANAKTTAKNALNLTSINNAQKALKSQIEGATTVAGVNO 997
QY 224 VVLAKEVEETLLASI-----DELATKAIGKKNNGLEANSKNTSLSGAYALSIDLI-- 276
DB 998 VSTTASSELNTAMSNLNGINDEAATAKAQKY-----TDADREKQTYNDNAVTAKTLLDK 1052


```
OY 277 -----AEKLVNK-----NEELKEKIDTAO-----CSTEFNKL 306
DB 1053 TAGSNDNKAVEOALORVNTAKTALNGDERLNEAKNTAKOQVATMSHLTDAOKANLTSQI 1112
OY 307 KSEHAVIGLNT-----TDDNAORAILKXKANKDGAELKLEKFAVENLSKAODTLKN 361
DB 1113 ESGTIVAGVIOAGNAGTLDOAMNOLROSIASND-ATKSESEYODANADLQNAVDAYTN 1171
OY 362 A 362
DB 1172 A 1172
```

RESULT 9

```
US-10-072-851-5835
; Sequence 5835, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072.851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 5835
; LENGTH: 2434
; TYPE: PRN
; ORGANISM: Staphylococcus aureus
US-10-072-851-5835
```

```
Query Match 10.1%; Score 182; DB 7; Length 2434;
Best Local Similarity 22.2%; Pred. No. 0.00057;
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;
```

```
OY 4 NNSGKDGNANSADSE---SVKGNPLTEISKITESNAV--VLAKE-VETLLASI----- 53
DB 702 NNNLEVAKQNAANTAIIDGLTSLNGPQAKLKEVGQATTLTPNVQTVRDNDQTLNTAKMGLR 761
OY 54 DELATKAIGKKNGLNGLANSQNTSLSGAYAISDLIAE---KLVNKLNEELKEKIDT 109
DB 762 DSIANEAT-IKAGQNTYDASQNKQTDYNSAVTAKAIIGQTSPSMAOEIQAQDOV-T 819
OY 110 AKOCSTFTNKLK-----SEHAVIGLNTLDDNAORAILK-----HANK-DKGAEELE 157
DB 820 AKQOALNGOENLRTAQTNAKQHLNGLSLDLD--AQKDAVKRQEGATHVEVYQAQNNAD 877
OY 158 KLEKAVENLSKAOD--TLK-----NAVKELTSP1--VHGNNSRKD-- 194
DB 878 ALNTAMTINLKNIGIQDQNTIKQGVNFTDADBAKRNAYTNNAVTOAEQILNKRAQGNPSKDV 937
OY 195 -----GNASTNSADESVKCP--NLTEISKITES-----NA 223
DB 938 ETALENVORAKNELNGNVAANKTAKNALNLTSINNAQKALASQIEGATTVAGVNO 997
OY 224 VVLAKEVETLLASI-----DELATKAIGKKNGLNGLANSQNTSLSGAYAISDLI-- 276
```

```
DB 998 VSTTASBELNAMSNIQNGINDEATATKAQY-----TDADREKQTAYNDAVTAKTLLDK 1052
OY 277 -----AEKLVNK-----NEELKEKIDTAO-----CSTEFNKL 306
DB 1053 TAGSNDNKAVEOALORVNTAKTALNGDERLNEAKNTAKOQVATMSHLTDAOKANLTSQI 1112
OY 307 KSEHAVIGLNT-----TDDNAORAILKXKANKDGAELKLEKFAVENLSKAODTLKN 361
DB 1113 ESGTIVAGVIOAGNAGTLDOAMNOLROSIASND-ATKSESEYODANADLQNAVDAYTN 1171
OY 362 A 362
DB 1172 A 1172
```

RESULT 10

```
US-09-815-242-12996
; Sequence 12996, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 12996
; LENGTH: 6281
; TYPE: PRN
; ORGANISM: Staphylococcus aureus
US-09-815-242-12996
```

```
Query Match 10.1%; Score 182; DB 6; Length 6281;
Best Local Similarity 22.2%; Pred. No. 0.0019;
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;
```

```
OY 4 NNSGKDGNANSADSE---SVKGNPLTEISKITESNAV--VLAKE-VETLLASI----- 53
DB 1481 NNNLEVAKQNAANTAIIDGLTSLNGPQAKLKEVGQATTLTPNVQTVRDNDQTLNTAKMGLR 1540
OY 54 DELATKAIGKKNGLNGLANSQNTSLSGAYAISDLIAE---KLVNKLNEELKEKIDT 109
DB 1541 DSIANEAT-IKAGQNTYDASQNKQTDYNSAVTAKAIIGQTSPSMAOEIQAQDOV-T 1598
OY 110 AKOCSTFTNKLK-----SEHAVIGLNTLDDNAORAILK-----HANK-DKGAEELE 157
DB 1599 AKQOALNGOENLRTAQTNAKQHLNGLSLDLD--AQKDAVKRQEGATHVEVYQAQNNAD 1656
OY 158 KLEKAVENLSKAOD--TLK-----NAVKELTSP1--VHGNNSRKD-- 194
```

```

Db 1657 ALNTAMTNLKNIGDONTIKOGVNFETDADAEKRNNAVTAQAEQILNKAAQPTSKDGV 1716
OY 195 -----GNASTNSADESVKCP--NLTEISKITTES-----NA 223
Db 1717 ETALENVORAKNELNGONAVNAKTTAKNALNNLTSTINNAQKEALKSQIEGATTVAGVNO 1776
OY 224 VLAFAVEETLLASI-----DELATKAIGKKGNGNGLEANOSKNTSLSGAVALSDLI-- 276
Db 1777 VSTTASELNTAMSNLQNGINDEAATKAQKY-----TDAREKOTAYNDAYTAANTLLDK 1831
OY 277 -----AEKLVNFK-----NEELKEKIDTAKQ-----CSTEFNTKL 306
Db 1832 TAGSNDNKAABEQAALQRYVNTAKTALNGDERLNEAKNTAKQOAVATWSHLTDAQKANLTSQI 1891
OY 307 KSEHAVLGLDNL-----TDDNAGRAILKKNKDKGAELEKLFKAVENLSKAADTLKN 361
Db 1892 ESGTTVAGVQIGIQAANAGTLDOAMNOLROSIAKSD--ATKSSSEDYODANADLQNAVDAYTN 1950
OY 362 A 362
1951 A 1951

```

RESULT 11
US-10-072-851-12996

Sequence 12996, Application US/10072851

```

GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits C
FILE REFERENCE: ELITRA 028A
CURRENT APPLICATION NUMBER: US/10/072.851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FASTSEQ for Windows Version 4.0
NO ID NO 12996
LENGTH: 6281
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-072-851-12996

```

Query Match 10.1%; Score 182; DB 7; Length 6281;
Best Local Similarity 22.2%; Pred. No. 0.0019;
Matches 107; Conservative 81; Mismatches 161; Indels 132; Gaps 24;

```

OY 4 NNSGKDNASANSAD--SVKGNLTFEISKRITESNAV--VLAVKE-VEETLLASI----- 53
Db 1481 NNNLEVAKNONANTALDITSLINGQOKAKLKVQOATTLTPVQVTRONAAOTLNTAMKGLR 1540
OY 54 DELATKAIGKKGNGNGLEANOSKNTSLSGAVALSDLIAB-----KLNVLNKEELKEKIDT 109
Db 1541 DSIINEAT-ITAGONYDASQNKQTDYNSAVTAKAKIIGOTSPSMAQELNKAQDQV-T 1598
OY 110 AKOCSTEFNTKLK-----SEHAVLGLDNLDDNAGRAILK-----HANK-DKGAEELE 157
Db 1599 AKQOALNGOENLRTAQNTAKOHLNGLSLDITD--AQKDAVAVROIEGATVHNEVTAQANNAD 1656

```

```

OY 158 KLFKAVENLSKAAD--TLK-----NAVKEITSP--VHGNSRKD-- 194
Db 1657 ALNTAMTNLKNIGDONTIKOGVNFETDADAEKRNNAVTAQAEQILNKAAQPTSKDGV 1716
OY 195 -----GNASTNSADESVKCP--NLTEISKITTES-----NA 223
Db 1717 ETALENVORAKNELNGONAVNAKTTAKNALNNLTSTINNAQKEALKSQIEGATTVAGVNO 1776
OY 224 VLAFAVEETLLASI-----DELATKAIGKKGNGNGLEANOSKNTSLSGAVALSDLI-- 276
Db 1777 VSTTASELNTAMSNLQNGINDEAATKAQKY-----TDAREKOTAYNDAYTAANTLLDK 1831
OY 277 -----AEKLVNFK-----NEELKEKIDTAKQ-----CSTEFNTKL 306
Db 1832 TAGSNDNKAABEQAALQRYVNTAKTALNGDERLNEAKNTAKQOAVATWSHLTDAQKANLTSQI 1891
OY 307 KSEHAVLGLDNL-----TDDNAGRAILKKNKDKGAELEKLFKAVENLSKAADTLKN 361
Db 1892 ESGTTVAGVQIGIQAANAGTLDOAMNOLROSIAKSD--ATKSSSEDYODANADLQNAVDAYTN 1950
OY 362 A 362
1951 A 1951

```

RESULT 12
US-10-032-585-7646

Sequence 7646, Application US/10032585

```

GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032.585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7646
LENGTH: 1881
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1881)..(1881)
OTHER INFORMATION: x=any amino acid
US-10-032-585-7646

```

Query Match 9.7%; Score 173.5; DB 7; Length 1881;
Best Local Similarity 25.6%; Pred. No. 0.0015;
Matches 112; Conservative 55; Mismatches 150; Indels 121; Gaps 25;

```

OY 2 ACNNSGRD-----GNASANSAD--ESVK-----GPNLTFEISKRITESNAVVLAVKE 45
Db 1470 ALTRKSSKDEVCNOKSELODSIKSVKSELKNFENKYNQETSLKMEIEKQ-----KE 1523
OY 46 VETLLASIDELAKKAIGKKGNGNGLEANOSKNTSLSGAVALSDLI--AEKLVNKEE 102
Db 1524 IYTLQTELKDRISSEVER-----AMLSENSETVIKRY--SDKIKSLSEKINSIKENH 1574
OY 103 LKEKIDTAKOCSTEFNTKLKSEHAVLGLDN-----LTDDNAG--RAILKKH----- 146
Db 1575 SKE-ITTHNEQKT-----SLKQDIATKLSQDHESAGQVLEDEKQKELKASLEKHNTESAT 1629
OY 147 --ANKDKGAELEKLFKAVENLSKAADTLKNAVVELTSPIVHGNSRKDNASTNS-- 201
Db 1630 SLEEKNNQIKELSETIKSLKTEIKTSGDALKOSQKEY-----KLTKTKNSDRESKLE 1681
OY 202 -----ADESVKGNLTFEISKRITESNAVVLAVKEVEETLLASIDELA 242
Db 1682 KOLELEKYSLOLQADERKLGITERETALK-----SELETYKNSGLSTTSELA 1730

```

243 TKAIGKKNNGLEANOSKNTSLSG--AYAISDL-----IAEKLNVLNKELEK--- 291
1731 --ALTKTV--KSLE-KEKEELOFLSGNKSKELEDYIOKHSIDSEKIKAL--TDELKEKTKQ 1784
292 IDTAOCSTEFNTKLKSEHAVALGLDNLTDNORAILKK-HANKDKGALEKLFKAVERN 350
1785 FDSCKKLTLELNDLTSTK-----KELETKQTSTKFKNLEERKDEIYKLNKELELLKN 1839
351 LSAAODTLKNAVKELTS 368
1840 DMSGAKKELSEKYSKLES 1857

RESULT 13
US-10-072-851-15590

Sequence 15590, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits C
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
NUMBER OF SEQ ID NOS: 1581
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15590
LENGTH: 1881
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1881)..
OTHER INFORMATION: X-any amino acid
US-10-072-851-15590

Query Match 9.7%; Score 173.5; DB 7; Length 1881;
Best Local Similarity 25.6%; Pred. No. 0.0015;
Matches 112; Conservative 55; Mismatches 150; Indels 121; Gaps 25;

2 ACNNSGKD-----GNASANSAD--ESVK-----GPNLTEISKITESNAVYLAKE 45
1470 ALTSSKDLLEVCGNQSKSELOLSKSVKSELKNFKENKYNQETSLADEIEKQ-----KE 1523
46 VETLASLDELATKAIGKKNNGLEANOSKNTSLSGAYAISDL---AEKLVNKLKNEE 102
1524 IVTLQTELEKDRSEVEKER-----AMLSENSETVIKEY--SDRIKSLESKINSIKEKH 1574
103 LKEKIDTAOCSTEFNTKLKSEHAVALGLDNLTDNNO-----RAILKKH----- 146
1575 SKF-ITTTNEOKT-----SLKODIAKLSQDHSQAQTOLEDEKQLEKELKSLKHNTHESAT 1629
147 --ANKDKGALEKLFKAVERNLSKAAODTLKNAVKELTSPIVHGNNSRKDGNASTNS--- 201
1630 SIEKNNQIKELSETIKSLKTELKSGDALKQSQKEY-----KTLTKKNSDSTRESKLE 1681
202 -----ADESVKGPNLTEISKITESNAVYLAKEVETL---LASIDELA 242

1682 KQLELEKYSKSDLOQTADEKLTQTEREIALK-----SELETVYVNSGLSTSELA 1730
243 TKAIGKKNNGLEANOSKNTSLSG--AYAISDL-----IAEKLNVLNKELEK--- 291
1731 --ALTKTV--KSLE-KEKEELOFLSGNKSKELEDYIOKHSIDSEKIKAL--TDELKEKTKQ 1784
292 IDTAOCSTEFNTKLKSEHAVALGLDNLTDNORAILKK-HANKDKGALEKLFKAVERN 350
1785 FDSCKKLTLELNDLTSTK-----KELETKQTSTKFKNLEERKDEIYKLNKELELLKN 1839
351 LSAAODTLKNAVKELTS 368
1840 DMSGAKKELSEKYSKLES 1857

RESULT 14
US-09-815-242-5815

Sequence 5815, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5815
LENGTH: 2076
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5815

Query Match 9.5%; Score 171.5; DB 6; Length 2076;
Best Local Similarity 20.0%; Pred. No. 0.0023;
Matches 85; Conservative 78; Mismatches 174; Indels 87; Gaps 14;

20 SVKGPNLTEISKITESNAVYL--AVKEVETLLASIDELATKA---IGKKNNGLEA-- 72
1132 TAKKQIIONNTATTEKQVVALNQVOEELATAINNINQADTNAEVDQAOLGTKAINAIO 1191
73 -----NOSKNTSL---SGAYAISDLIAEKLNVLNKELEKIDTAOCSTE 116
1192 PNIVKPAALAOINQHYNAKLAETINATPDATNDEKNAATVTL--NODROQATIESIKQAN-- 1248
117 FNNKLESEHAVALGLDNLTDNORAILKKHANKDKGALEKLFKAVERNLSKAAODTLKN 176
1249 -TNAEVDQAATYAENNI--DAVGVDVYVKQQAARDKITAELVAKRIEAVKQPNATDEKQA 1305
177 AVKELTS-----PIVHGNNSRKDGNASTN-----SADESVKGPNLTEISKITES 221

```

Db 1306 AVNQINOLKDAQAINQONQNDQVDTTNOAVNAIDNVEAEVYIKPRAIDIEKAVK 1365
QY 222 N-----AVVLAVKEVEETLLASIDELATKAIGKIGNNGLEANSKNTSL 265
Db 1366 QOQIDNSLDSTDNEKEVASQALAKEKEKALAIIDQAQINSQVNAATNGVSA-----IKI 1420
QY 266 LSGAVAISDLIAEKLNVKNELEKEKIDTAQCS-----TEFTNKLKSEHAVLG 314
Db 1421 IQPETHKVPAREKIN-QKANELRAKINQDKKATAEERQVALDKINEFVNQAMTDITNKR 1479
QY 315 LDNLTDNAQRAI-----LKKHANKDKGAAELEKLEKFAVENLSKAADTLKNAVK 364
Db 1480 TNOQVDDTTSQALDSIALVTPDHIYRAAADVAQVKKREIEQAHAHATDEEKQVALN 1539
QY 365 ELTS 368
Db 1540 QLAN 1543

```

```

BLT 15
Sequence 5815, Application US/10072851
GENERAL INFORMATION:

```

```

APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits C
FILE REFERENCE: EITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ. ID NOS: 15811
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5815
LENGTH: 2076
TYPE: PRT
ORGANISM: Staphylococcus aureus
US/10-072-851-5815

```

```

Query Match 9.5%; Score 171.5; DB 7; Length 2076;
Best Local Similarity 20.0%; Pred. No. 0.0023;
Matches 85; Conservative 78; Mismatches 174; Indels 87; Gaps 14;

```

```

QY 20 SVKGPNLTEISKTTESNAVVL-AVKEVEETLLASIDELATKA---IGKIGNNGLEA-- 72
Db 1132 TAKQOEINQNTNATTEERQVALNOVDDELATANNINQADTNAEYDQAQOAGTKAINAIQ 1191
QY 73 -----NOSKNTSL--SGAVAISDLIAEKLNVKNELEKEKIDTAQCS 116
Db 1192 PNLYKKPAALAIQINOHNAKLAEINATPDATNDEKNAIINTL-NODRQQALESITQAN-- 1248
QY 117 FTNKLKSEHAVLGDNLTDDNAQRAILKHKANKDKGAAELEKLEKFAVENLSKAADTLKN 176
Db 1249 -TNAEVDQAATVAENNI--DAVQYDVYVKKQAARDKITAFAVAKRTEAVKQTPNATDEEKQA 1305
QY 177 AVKEELTS-----PIVGNNSRKDKGNASTN-----SADESYKGPNLTEISKTTES 221
Db 1306 AVNQINOLKDAQAINQONQNDQVDTTNOAVNAIDNVEAEVYIKPRAIDIEKAVK 1365

```

```

QY 222 N-----AVVLAVKEVEETLLASIDELATKAIGKIGNNGLEANSKNTSL 265
Db 1366 QOQIDNSLDSTDNEKEVASQALAKEKEKALAIIDQAQINSQVNAATNGVSA-----IKI 1420
QY 266 LSGAVAISDLIAEKLNVKNELEKEKIDTAQCS-----TEFTNKLKSEHAVLG 314
Db 1421 IQPETHKVPAREKIN-QKANELRAKINQDKKATAEERQVALDKINEFVNQAMTDITNKR 1479
QY 315 LDNLTDNAQRAI-----LKKHANKDKGAAELEKLEKFAVENLSKAADTLKNAVK 364
Db 1480 TNOQVDDTTSQALDSIALVTPDHIYRAAADVAQVKKREIEQAHAHATDEEKQVALN 1539
QY 365 ELTS 368
Db 1540 QLAN 1543

```

```

Search completed: March 18, 2002, 09:58:28
Job time: 386 sec

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:48 ; Search time 68.77 Seconds
(without alignments)
407.623 Million cell updates/sec

Title: US-09-596-746a-38

Perfect score: 1797

Sequence: 1 MACNNSGKDGMSANSADSES.....ENLSKAQDTLKNVKEITS 368

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	50.1	209	2 I40142	outer surface prot
2	661	36.8	193	2 S70287	outer surface prot
3	646	35.9	211	2 I40145	outer surface prot
4	642	35.7	193	2 S70279	outer surface prot
5	639	35.6	191	2 I40153	outer surface prot
6	623.5	34.7	210	2 G70218	outer surface prot
7	622	34.6	209	2 S69917	outer surface prot
8	619.5	34.5	210	2 S69927	outer surface prot
9	619	34.4	193	2 S70286	outer surface prot
10	614	34.2	191	2 S70284	outer surface prot
11	612.5	34.1	212	2 S69922	outer surface prot
12	605.5	33.7	194	2 S70268	outer surface prot
13	605	33.7	178	2 I40104	outer surface prot
14	603.5	33.7	178	2 I40125	outer surface prot
15	603.5	33.6	210	2 I40144	outer surface prot
16	596.5	33.2	212	2 I40143	outer surface prot
17	589	32.8	193	2 S70274	outer surface prot
18	588.5	32.7	212	2 S20543	outer surface prot
19	588	32.7	211	2 I40268	outer surface prot
20	587	32.7	211	2 S69918	outer surface prot
21	586	32.6	211	2 S69932	outer surface prot
22	584.5	32.5	212	2 S70254	outer surface prot
23	581	32.3	193	2 S70280	outer surface prot
24	579.5	32.2	212	2 S69921	outer surface prot
25	577	32.1	211	2 I40278	outer surface prot
26	575	32.0	193	2 S70265	outer surface prot
27	572	31.8	193	2 S70276	outer surface prot
28	569.5	31.7	177	2 I40129	outer surface prot
29	569	31.7	207	2 I40271	outer surface prot

30	567	31.6	200	2 I40122	outer surface prot
31	567	31.6	207	2 S69919	outer surface prot
32	567	31.6	207	2 S37727	outer surface prot
33	566.5	31.5	190	2 S70261	outer surface prot
34	566.5	31.5	194	2 S70277	outer surface prot
35	566.5	31.5	212	2 I40279	outer surface prot
36	566	31.5	209	2 I40281	outer surface prot
37	565.5	31.5	194	2 S70289	outer surface prot
38	565	31.4	211	2 I40277	outer surface prot
39	564	31.4	207	2 I40276	outer surface prot
40	562.5	31.3	214	2 S69916	outer surface prot
41	562	31.3	209	2 I40285	outer surface prot
42	560.5	31.2	180	2 I40110	outer surface prot
43	559	31.1	189	2 S70263	outer surface prot
44	559	31.1	203	2 I40108	outer surface prot
45	558	31.1	178	2 I40124	outer surface prot

ALIGNMENTS

```
RESULT 1
I40142
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40142; S70283
R:Stevenson, B.; Bockenstedt, L.K.; Barthold, S.W.
Infect. Immun. 62, 3568-3571, 1994
A:Title: Expression and gene sequence of outer surface protein C of Borrelia burgdorferi
A:Reference number: I40142; MUID:94314484
A:Accession: I40142
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-209 <RES>
A:Cross-references: EMBL:U04240; NID:g2314877; PIDN:AAC45538.1; PID:g434658
R:Liway, I.; Gibbs, C.P.; Schuster, R.; Dornet, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease
A:Reference number: S70255; MUID:96296448
A:Accession: S70283
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-209 <LIV>
A:Cross-references: EMBL:L42894; NID:g858722; PIDN:AAB37002.1; PID:g1695219
A:Experimental source: strain 28691
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 50.1%; Score 901; DB 2; Length 209;
Best Local Similarity 98.9%; Pred. No. 2, 2e-35;
Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNNSGKDGMSANSADSEVKPNTETSKTTESNAVYLVAKVEYTLASIDELATKA 60
:::|||||
DB 17 ISGNNGKDGMSANSADSEVKPNTETSKTTESNAVYLVAKVEYTLASIDELATKA 76
QY 61 IGGKINNGLEANSKNTSLSGAVYISDLIAKLVNKKKEEIKETIDTAOKOSTEPTNK 120
|||||
DB 77 IGGKINNGLEANSKNTSLSGAVYISDLIAKLVNKKKEEIKETIDTAOKOSTEPTNK 136
QY 121 LKSEHAVLGLDNLTDNAQRAILKKHANKDKGAEELEKLFKAVENTSKAQDTLKNVKE 180
|||||
DB 137 LKSEHAVLGLDNLTDNAQRAILKKHANKDKGAEELEKLFKAVENTSKAQDTLKNVKE 196
QY 181 LKSPYV 186
|||||
DB 197 LKSPYV 202
RESULT 2
```

S70287

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000

C:Accession: S70287

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A:Reference number: S70255; MUID:96296448

A:Accession: S70287

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Cross-references: EMBL:142895; NID:9858723; PID:AA37003.1; PID:g1695220

A:Experimental source: strain 28354

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 36.8%; Score 661; DB 2; Length 193;

Matches 139; Conservative 20; Mismatches 25; Indels 2; Gaps 2;

QY 3 CNGSGKDGNSANSADSVKGNLTETSKRTESNAVLAKEVEYLASIDELATKAI 62

Db 1 CNGSGKDGNTSANSADSVKGNLTETSKRTESNAVLAKEVEYLASIDELATKAI 60

QY 63 KRI-GNNGLEANOSKNTSLSGAVASDLIAEKLNVKLN-EELKEKIDPAKOCSTEF 120

Db 61 KRIQONGGLAVEAGHNGTLAGAVYITSLITQKLDGLKNSKEKEIKENAKKCSDEFTK 120

QY 121 LKSEHAVLGIDNLTDDNAQRALIKKANKDKGAELKFKAVENISKAADTLKNAVE 180

Db 121 LKSEHAVLGIDNLTDDNAQRALIKKANKDKGAELKFKAVENISKAADTLKNAVE 180

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 181 LTSPIV 186

Db 181 LTSPIV 186

QY 119 NKIKSEHAVLGIDNLTDDNAQRALIKKANKDKGAELKFKAVENISKAADTLKNAV 178

Db 137 NKIKSSHTTELKQDODDADKAKALLTHTNKDKGAELDKFKAVENISKAADTLKNAV 196

QY 179 KELTSPIV 186

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204

Db 197 KELTSPIV 204

QY 197 KELTSPIV 204


```

Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 1 MACNNSGKDGNASANSADSEYKGPMLTEISKKITBESNAVLAWEVEFLLASIDELATKA 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 17 ISCNNSGKDGMTSANSADSEYKGPMLTEISKKITDSNNAVLAWEVEFLLASIDELATKA 76

QY 61 IGKRI-GNNGLEANOSKRTSLISGAYAIISDLIAEKLNVLKKNELEKIDTPKOCSTEF 119
      ||||| ||||| : |||:||||| || :||: ||||| ||||| ||||| ||||| |||
Db 77 IGKRIHQNGNDLTENNHGSLAGAYAIISLTIKOKLDGLKNEGLEKIKIDAAKCSSEFTN 136

QY 120 KLESEHAVALGDLNDDNAQRALIKKHANKDKGAELKELFRAVENLSKAADPTLKNAVK 179
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 137 KLEKHTDGLGEGYTDADAKRAILIKTNTKTKGAELCKLESEVLSKAKEMLANSYK 196

QY 180 ELTSPIV 186
      |||||:|
Db 197 LTSPV 203

RESULT 7
569917
outer surface protein C precursor - Lyme disease spirochete (strain B. pacificus)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
A:Variety: strain B. pacificus
C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
A:Accession: S69917; S72679
R:Jauris-Heipke, S.; Lileg, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek,
J. Clin. Microbiol. 33, 1860-1866, 1995
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A:Reference number: 140047; MUID:95395018
A:Accession: S69917
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-209 <JAU>
A:Cross-references: EMBL:X83555
A:Experimental source: strain B. pacificus
R:Roessler, D
submitted to the EMBL Data Library, January 1995
A:Reference number: S72679
A:Accession: S72679
A:Molecule type: DNA
A:Residues: 1-50, 'E', 52-61, 'I', 63-117, 'D', 119-122, 'V', 124-173, 'D', 175-176, 'D', 178-209
A:Cross-references: EMBL:X83555; NID:9872019; PIDN:CA58545.1; PID:9872020
C:Superfamily: Lyme disease spirochete surface protein C
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-209/Product: outer surface protein C #status predicted <MAT>

Query Match 34.6%; Score 622; DB 2; Length 209;
Best Local Similarity 69.9%; Pred. No. 1.6e-22;
Matches 130; Conservative 24; Mismatches 32; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNASANSADSEYKGPMLTEISKKITBESNAVLAWEVEFLLASIDELATKA 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 17 ISCNNSGKDGMTSANSADSEYKGPMLTEISKKITDSNNAVLAWEVEFLLASIDELATKA 76

QY 61 IGKRIINGGLEANOSKRTSLISGAYAIISDLIAEKLNVLKKNELEKIDTANOCSTEF 120
      ||||| ||||| :|||:||||| || :||: ||||| ||||| ||||| ||||| |||
Db 77 IGKRIADANGLGVAQNQNSLLAGAYAIISLTIKOKLSALNSELEKIKAVKCSSEDP 136

QY 121 LKSEHAVALGDLNDDNAQRALIKKHANKDKGAELKELFRAVENLSKAADPTLKNAVK 180
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 137 LKNGNAQGLAATDDNKAALIKLTNGTNDGAKELKELSEVESLVRAQAQMLTNSYK 196

QY 181 LTSPIV 186
      |||||:|
Db 197 LTSPV 202

RESULT 8
569927
outer surface protein C precursor - Lyme disease spirochete (strain Pk)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

```

Query Match	34.58;	Score	619.5;	DB	2;	Length	210;
Best Local Similarity	69.08;	Pred. No.	2.1e-22;				
Matches	129;	Conservative	25;	Mismatches	32;	Indels	1;
						Gaps	1;

1869

Query Match	34.48;	Score 619;	DB 2;	Length 193;
Best Local Similarity	71.18;	Pred. No. 2e-22;		
Matches 133;	Conservative 22;	Mismatches 28;	Indels 4;	Gaps 3

QY 3 C NNSGKDGNASANSDESVKGPNTLISKITESSNAVVLAKVEVTLASIDELATKAIG 62
| | | | | | | | | | | | | | | | | : | | | | |
d

RESULT 10
670394

Query Match	34.28;	Score 614;	DB 2;	Length 191;
Best Local Similarity	69.68;	Pred. No. 3.3e-22;		
Matches 128; Conservative	22;	Mismatches 34;	Indels 0;	Gaps 0

[illegible]

```

RESULT 11
569922
outer surface protein C precursor - Borrelia afzelii (strain PLj7)
C:Species: Borrelia afzelii
A:Variety: strain PLj7
C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C:Accession: S69922; S72675
R:Jarris-Helpke, S.; Lilel, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A:Reference number: I40047; M01D:95395018
A:Accession: S69922
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-212 <TANU>
A:Cross-references: EMBL:X01523

```

A: Experimental source: strain PLj7
R: Roessler, D.
submitted to the EMBL Data Library, September 1994
A: Reference number: S72674
A: Accession: S72675
A: Molecule type: DNA
A: Residues: 1-125, 'V', 127-154, 'E', 156-212 <ROE>
A: Cross-references: EMBL:X81523; NID:g804956; PIDN:CAAS7243.1; PID:g804957
C: Genetics:
A: Gene: ospC
C: Superfamily: Lyme disease spirochete surface protein C

Query Match	34.1%	Score 612.5	DB 2	Length 212;
Best Local Similarity	69.3%	Pred. No. 4,4e+22;		
Matches 131; Conservative	27;	Mismatches 26;		
			Indels 3;	Gaps 3;

[illegible]

RESULT 12
 S70268
 Outer surface protein C - Lyme disease spirochete
 C.Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C.Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C.Accession: S70268
 R.Lacey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A.Title: Evidence for lateral transfer and recombination in Lyme disease
 A.Reference number: S70255; MUID:96296448
 A.Accession: S70268
 A.Status: nucleic acid sequence not shown
 A.Molecule type: DNA
 A.Residues: 1-194 <LIV>
 A.Cross-references: EMBL:L42888; NID:g858716; PIDN:AAB36996.1; PID:g1695213
 Experimental source: strain H9
 A.Notes:
 A.Organism: spirochete
 A.Superfamily: Lyme disease spirochete surface protein C

Query Match	33.7%	Score 605.5	DB 2	Length 194
Best Local Similarity	69.0%	Pred No. 8.3e-22		
Matches 129	Conservative 27	Mismatches 28	Indels 3	Gaps 3

```
0Y      3 C NNSCKDC - MANSANGADSVGPNLTELISKTTESNAVLAVKVEFTLLASIDELAKAI   61
        | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 C NNSCGDSDASTNPDADESAGPNLTEISKRITTSMAFVLAVKEVETLVSSIDELAQAII   60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y      62 GKKIKNNGLEAOSKNTSLSGAVAISDLIAEKLNVL - KNE - ELKEKIDTAOKOSTEFTN 119
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 GKRIQNNBELAEQONONGSILGAAVAISALIRTKKIDELTKNGBELGEYEKAKKCSEFTN 120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y      120 KIKSEHAVIGLDNLTDNOAQARALLKKHANRKGAAELEKTLKFAVENLSKAQODTLKANVK 179
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 KIKGGHAEGLAAAAADENAKRAILLKTNTQTKDKGAEBELEKLFKSVESSLAKAAKRESLTNSVK 180
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y      180 ELTSPIV 186
        | | | | | | | | | |
```

Db 181 EETNPV 187

RESULT 13

140104

outer surface protein C - Lyme disease spirochete (strain 272) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000

C:Accession: 140104; S54184

R:Bacterion, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

J:Publication: 177, 3036-3044, 1995

A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A:Reference number: 140104; MUID:95286481

A:Accession: 140104

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178 <RES>

A:Cross-references: EMBL:X84785; NID:g793787; PIDN:CAA59256.1; PID:g793788

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match	33.78%	Score 605,	DB 2:	length 176:
Best Local Similarity	73.08%	Pred. No.	7, 9e-22:	
Matches 130,	Conservative 19,	Mismatches 27,	Indels 2,	Gaps 2:

```
QY      11 NASASAEESVKGPLUTISKKITFTSNVAVLAVEVEFLLASIDELATKAGIKRT - GNGG   69
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 NTSASASAEESVKGPLUTISKKITFTSNVAVLAVEVEFLLASIDELATKAGIKRI QONGG   60

QY      70 LEANOSKMTSLSGAYASIDIEAETKLNVLNKN-ELEKEKERIDTAKOCSTFPTTKLKSEHAVL  122
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      61 LAVEGHHGTLLAGAYTTISKLTQKLDCLKSKSEKITETENNAKRCSDEPTFKLBEBHQD  122

QY      129 GUDNLTDDNOARPAIIKTHANKNDGAAELEKEJLFKAVENTSKAAODTLNAKVVELTSPIV  186
        | : | : | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 GIENVTDENAKKAILITDAADKGAEELEKEJLFKAVENTLPKAAKEHLANSVKEILTSPIV  178
```

RESULT 14
140125
outer surface protein C - Lyme disease spirochete (strain MUL) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40125; S54195
R:Thielsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebesch, A.M.; Hansen, K
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; MUID:95286481
A:Accession: I40125
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL.X84779; NID:g793817; PIDN:CA59250.1; PID:g793818
C:Superfamily: Lyme disease spirochete surface protein C

Query Match	33.78;	Score 605;	DB 2;	Length 178;
Best Local Similarity	73.08;	Pred. No. 7.9e-22;		
Matches 130; Conservative	19;	Mismatches 27;	Indels 2;	Gaps 2

[illegible]

RESULT 15

140144

outer surface protein C precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000

C:Accession: 140144; S70282

R:Stevenson, B.; Barthold, S.W.

FEMS Microbiol. Lett. 124, 367-372, 1994

A:Title: Expression and sequence of outer surface protein C among North American isolate

A:Reference number: 140143; MUID:95154673

A:Accession: 140144

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-210 <RES>

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664

Search completed: March 18, 2002, 09:56:48
 Job time: 371 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:52 ; Search time 39.62 Seconds
(without alignments)
340.552 Million cell updates/sec

Title: US-09-596-746A-38
Perfect score: 1797

Sequence: 1 MACNNSGKQDNASNADES.....ENLSKAQDTLKNAVKELTS 368

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the entry being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	623.5	34.7	210	1	OSCL_BORBU
2	588.5	32.7	212	1	OSCL_BORBU
3	346	19.3	214	1	VM24_BORHE
4	311.5	17.3	215	1	VM03_BORHE
5	172.5	9.6	1790	1	USO1_YEAST
6	166.5	9.3	944	1	NUP1_YEAST
7	163.5	9.1	1167	1	CAGA_HELPY
8	162.5	9.0	2017	1	MYSN_DROME
9	159.5	8.9	1875	1	MLP1_YEAST
10	158.5	8.8	705	1	YNP9_CAEEL
11	157.5	8.8	1251	1	RBP2_PLAVB
12	157	8.7	1639	1	LMG1_DROME
13	156.5	8.7	1013	1	AGOD_DROME
14	154	8.6	1526	1	MYSD_SCHPO
15	153	8.5	548	1	CEAK_ECOLI
16	152.5	8.5	1391	1	MST2_DROHY
17	151.5	8.4	1966	1	MYSB_CAEEL
18	150	8.3	1147	1	CGAL_HELPY
19	149.5	8.3	1786	1	LMB1_HUMAN
20	149.5	8.3	2116	1	MYSD_DICDI
21	148.5	8.3	1957	1	YDB6_SCHPO
22	148.5	8.3	2249	1	TPR_HUMAN
23	148	8.2	1433	1	REST_CHICK
24	148	8.2	1701	1	MSP1_PLAFR
25	147.5	8.2	1102	1	MYSC_CHICK
26	147.5	8.2	1182	1	CGA2_HELPY
27	147	8.2	845	1	SCP1_MESAU
28	147	8.2	918	1	YMOB_CAEEL
29	147	8.2	1427	1	REST_HUMAN
30	146.5	8.2	1637	1	MSP1_PLAFR
31	146.5	8.2	1786	1	LMB1_MOUSE
32	145.5	8.1	1938	1	MYH4_RABIT
33	145	8.1	539	1	MYH3_HYDAT

34	144.5	8.0	1044	1	YAF3_SCHPO	00857 schizosacch
35	144	8.0	483	1	M6_STRPY	P08089 streptococ
36	143.5	8.0	1509	1	MYSN_ACACA	P05659 acanthamoeb
37	143	8.0	364	1	VM21_BORHE	P21875 borrelia he
38	143	8.0	1186	1	CAGA_HELPY	P55980 helicobacte
39	143	8.0	1727	1	ALM1_SCHPO	09ute5 schizosacch
40	142.5	7.9	2230	1	GOG4_HUMAN	013439 homo sapien
41	142	7.9	522	1	CEAL_HUMAN	P02978 escherichia
42	142	7.9	1186	1	SMC_BACSU	P51834 bacillus su
43	141.5	7.9	564	1	M12_STRPY	P19401 streptococ
44	141	7.8	276	1	SFAS_CHLRE	039618 chlamydomon
45	141	7.8	1701	1	MSP1_PLAFM	P08569 plasmodium

ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD	PRT	210 AA.
AC	007337				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).				
GN	OSPC OR BB819.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=9329332; PubMed=8098841;				
RA	Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,				
RA	Will G., Wilske B.;				
RT	"Genetic heterogeneity of the genes coding for the outer surface				
RT	protein C (Ospc) and the flagellin of Borrelia burgdorferi.";				
RL	Med. Microbiol. Immunol. 182:37-50(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=9329332; PubMed=8478108;				
RA	Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,				
RA	Schwab E., Wanner G.;				
RT	"Immunological and molecular polymorphisms of OspC, an immunodominant				
RT	major outer surface protein of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:2182-2191(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=94041630; PubMed=8225587;				
RA	Padua S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;				
RT	"Molecular characterization and expression of p23 (OspC) from a North				
RT	American strain of Borrelia burgdorferi.";				
RL	Infect. Immun. 61:5097-5105(1993).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=96025162; PubMed=7494039;				
RA	Fukunaga M., Hamase A.;				
RT	"Outer surface protein C gene sequence analysis of Borrelia				
RT	burgdorferi sensu lato isolates from Japan.";				
RL	J. Clin. Microbiol. 33:2415-2420(1995).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE=98065943; PubMed=9403685;				
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,				
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,				
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,				
RA	Peterson J., Kellavag A.R., Quackenbush J., Salzberg S., Hanson M.,				
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,				

RA Uterback T., Matthey L., McDonald L., Artlach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, *Borrelia*
 RT burgdorferi";
 RL Nature 350:580-586(1997).
 CC - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
 CC - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69596; CAA49306.1; -;
 DR EMBL: U01894; AAA16058.1; -;
 DR EMBL: D49497; BAA08457.1; -;
 DR EMBL: AE000792; AAC66329.1; -;
 DR TIGR: BB19; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF001441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
 KW Complete proteome.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 210 OUTER SURFACE PROTEIN C.
 FT LIPID 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
 SQ SEQUENCE 210 AA: 22340 MW: 744FC97BF91777BF CRC64;

Query Match 34.7%; Score 623.5; DB 1; Length 210;
 Best Local Similarity 69.5%; Pred. No. 5.3e-22;
 Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 1 MACNNSGKDGMSANSADSVKGNPLTEISKITTESNAVAVLVKVEVETLLASIDELATKA 60
 DB 17 ISCNNSGKGNTSANSADSVKGNPLTEISKITTSNNAVAVLVKVEVETLLASIDELATKA 76
 QY 61 ICKKI-GNNGLEANSKNTSLSGAIAISDLIAEKLNVKNEKEKIDTAQCSSTFTN 119
 DB 77 ICKKHONNGDLTENNNHSLLAGAIAISTLIKOKLDIGKNGLEKEKIDAAKCSSTFTN 136
 DB 120 KLSSEHAVLGLDNLTDNNAORAILKKHANKDKGAELKFRVAVENLSKAODITLKNV 179
 DB 137 KLEKHTDLGKRGVTDADAKKAILKTNGTKTGAELEKLEFSEVYLSKAKKEMLVANSV 196
 QY 180 ELTSPV 186
 DB 197 ELTSPV 203

RESULT 2
 OSC2_BORBU STANDARD; PRT; 212 AA.
 AC Q08137;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR (PC).
 GN OSPC.
 OS *Borrelia burgdorferi* (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=PRO;
 RX MEDLINE=9221995; PubMed=1560779;
 RA Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

RA Soutschek E.;
 RT "Molecular analysis and expression of a *Borrelia burgdorferi* gene
 RT encoding a 22 kDa protein (PC) in *Escherichia coli*.";
 RL Mol. Microbiol. 6:503-509(1992).
 RN [2]
 RP SEQUENCE OF 1-205 FROM N.A.
 RC STRAIN=DK26;
 RA MEDLINE=94075528; PubMed=8253951;
 RX Theisen M., Frederixsen B., Lebech A.M., Vuust J., Hansen K.;
 RT "Polymorphism in *ospc* gene of *Borrelia burgdorferi* and
 RT immunoreactivity of *ospc* protein: implications for taxonomy and for
 RT use of *ospc* protein as a diagnostic antigen.";
 RL J. Clin. Microbiol. 31:2570-2576(1993).
 CC - FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
 CC - SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62162; CAA44093.1; -;
 DR EMBL: X73624; CAA52003.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF001441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 212 OUTER SURFACE PROTEIN C.
 FT LIPID 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
 SQ SEQUENCE 212 AA: 22499 MW: C206C231FBE2E7D4 CRC64;

Query Match 32.7%; Score 588.5; DB 1; Length 212;
 Best Local Similarity 68.8%; Pred. No. 1.9e-20;
 Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNNSGKDGMSANSADSVKGNPLTEISKITTESNAVAVLVKVEVETLLASIDELATKA 59
 DB 17 ISCNNSGKGDSASTNPADSAKGNPLTEISKITTSNNAVAVLVKVEVETLLASIDELATKA 76
 QY 60 AIGKKI-GNNGLEANSKNTSLSGAIAISDLIAEKLNVKN-BELKEKIDTAQCSSTFTN 117
 DB 77 AIGKIDNNNGGLAALNNNGSLLAGAIAISTLYTEKLSKLEKEKIDTAQAKKCSSTFTN 136
 QY 118 TNKLSSEHAVLGLDNLTDNNAORAILKKHANKDKGAELKFRVAVENLSKAODITLKNV 177
 DB 137 TNKLSGSHADLGKQDADDDAKKAILKTHTTKGAKERFDLFSEVGLKKAQVALTNS 196
 QY 178 VKELTSPV 186
 DB 197 VKELTSPV 205

RESULT 3
 VM24_BORHE STANDARD; PRT; 214 AA.
 AC P32778;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
 GN VMP24.
 OS *Borrelia hermslii*.
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.
 OX NCBI_TaxID=140;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-SSP, HSI SEROTYPE 24;
 RX MEDLINE-9313110; PubMed-1484486;
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
 RT "Subtelomeric expression regions of *Borrelia hermslii* linear plasmids
 are highly polymorphic."
 CC Mol. Microbiol. 6:3299-3311(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L04786; AAA22964.1; -
 CC InterPro: IPR001800; Lipoprotein_6.
 CC Pfam: PF01441; Lipoprotein_6; 1.
 CC ProDom: PD001149; Lipoprotein_6; 1.
 CC DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC KW Outer membrane; Lipoprotein; signal; plasmid.
 CC FT SIGNAL 1 18 PROBABLE.
 CC FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE
 CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 CC FT SEQUENCE 214 AA; 22541 MW; F1583F510246FC7 CRC64;

Query Match 19.3%; Score 346; DB 1; Length 214;
 Best Local Similarity 42.1%; Pred. No. 1.2e-09;
 Matches 82; Conservative 40; Mismatches 57; Indels 16; Gaps 5;

OY 1 MACNSGKQGNASANSDESVKGP-----NLTETSKKTTESNAVAVLAKVEVETLLASIDEL 56
 DB 17 MSCNNGPPE-----LKSDEVAKSDGTVLDLAKYSKKIKKSAFSAVKEVETLVKSYDEL 71
 OY 57 ATKAIGKTKIGN-NLEPANSKNTSLSGAFAISDLAEKLVK-----NEELKEKIDTA 110
 DB 72 A-KATKRIKKNBGLDTBAGQNSLTAQVHSYSAVKIKGALETTSGISNELKTKITEV 130
 OY 111 KOCSTFTKTLKSEHVAVLGLDNLTDNAORAILKHKANKDKGAELKFEKAVENISKAA 170
 DB 131 KSKAEAFLLNKDKGHELGKKDASDDTKKAIKKKNSDKTGGASLEALNTAVDALILKAA 190
 OY 171 QDTLKNVKELTSP 185
 DB 191 EGEVEAIRELTAPV 205

RESULT 4
 VMO3_BORHE STANDARD; PRT; 215 AA.
 ID VMO3_BORHE
 AC Q02448;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
 GN VMP3.
 OS *Borrelia hermslii*.
 OC Plasmid.
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.
 OX NCBI_Taxid=140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SSP, HSI SEROTYPE 3;
 RX MEDLINE-9313110; PubMed-1484486;
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
 RT "Subtelomeric expression regions of *Borrelia hermslii* linear plasmids

RT are highly polymorphic."
 RL Mol. Microbiol. 6:3299-3311(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP24.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L04789; AAA22967.1; -
 CC InterPro: IPR001800; Lipoprotein_6.
 CC Pfam: PF01441; Lipoprotein_6; 1.
 CC ProDom: PD001149; Lipoprotein_6; 1.
 CC DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC KW Outer membrane; Lipoprotein; signal; plasmid.
 CC FT SIGNAL 1 18 PROBABLE.
 CC FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
 CC FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 CC FT SEQUENCE 215 AA; 23139 MW; 684C74D5F87C771 CRC64;

Query Match 17.3%; Score 311.5; DB 1; Length 215;
 Best Local Similarity 39.8%; Pred. No. 4e-08;
 Matches 78; Conservative 36; Mismatches 65; Indels 17; Gaps 5;

OY 1 MACNSGKQGNASANSDESVKGP-----NLTETSKKTTESNAVAVLAKVEVETLLASIDEL 56
 DB 17 MSCNNGPPE-----LKSDEVAKSDGTVLDLAKYSKKIKKSAFSAVKEVETLVKSYDEL 71
 OY 57 ATKAIGKTKIGN-NLEPANSKNTSLSGAFAISDLAEKLVK-----NEELKEKIDT 109
 DB 72 A-KATKRIKKNBGLDTBAGQNSLTAQVHSYSAVKIKGALETTSGISNELKTKITEV 130
 OY 110 AKOCSTFTKTLKSEHVAVLGLDNLTDNAORAILKHKANKDKGAELKFEKAVENISKAA 169
 DB 131 KKESEAFVQVSKKRIIDLKEGVTDAAKASAILVDTGDKGAELIKLTAIDELKAA 190
 OY 170 AODTLKNVKELTSP 185
 DB 191 ANDAVETVIRELTASV 206

RESULT 5
 USOL_YEAST STANDARD; PRT; 1790 AA.
 ID USOL_YEAST
 AC P25386;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL1.
 GN USOL1 OR INT1 OR YDL058W.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;
 OC *Saccharomycetaceae*; *Saccharomycetaceae*; *Saccharomyces*.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-X2180-1A.
 RX MEDLINE-91185402; PubMed-2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 Yamasaki M.;
 RT "A cytoskeleton-related gene, usol1, is required for intracellular
 RT protein transport in *Saccharomyces cerevisiae*."
 RL J. Cell Biol. 113:245-260(1991).
 RN [2]

RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
 RA Kendrick K.E.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Bai Y., Symington L.S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
 CC ER AND THE GOLGI COMPLEX.
 CC -1- DOMAIN: THE RODLIKE TAIL. SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
 CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 DR EMBL: X54378; CA38253.1; -
 DR EMBL: L03188; AAB00143.1; -
 DR EMBL: U53668; AAB6659.1; -
 DR PIR: A38455; A38455.
 DR HSP: P80220; IDP.
 DR SGD: S0002216; USOL.
 DR InterPro: IPR002017; Spectrin.
 KM Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
 FT DOMAIN 1 724
 FT DOMAIN 725 1790
 FT DOMAIN 465 487
 FT DOMAIN 991 1790
 FT DOMAIN 1172 1786
 FT CONFLICT 847 847
 FT CONFLICT 924 924
 FT CONFLICT 1253 1253
 FT CONFLICT 1319 1319
 FT CONFLICT 1461 1461
 FT CONFLICT 1581 1581
 FT CONFLICT 1600 1600
 FT CONFLICT 1661 1661
 FT CONFLICT 1772 1772
 SC SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;
 Query Match 9.6%; Score 172.5; DB 1; Length 1790;
 Best Local Similarity 22.1%; Pred. No. 0.6;
 Matches 103; Conservative 66; Mismatches 152; Indels 145; Gaps 20;
 QY 3 CUNSCRGD-----NASANSDESIVKPNLFEISKI-----TESNAVVLAVK 44
 DB 929 CUNLSKEKEHISKELVEKSRFGSHDNV--AKLEFKLSANNYKQMAENESLRAVE 986
 QY 45 E-----VETLASIDELATKAIGKIGNGLEAN--OSKNTSLSGAVASIDLAEK 94
 DB 987 ESKNESSIQLSNLONKIDMSOEKNEFCIEGSTEKNIEQLKKT-----ISDLQTK 1038
 QY 95 LNVLKNEB-----LKEKIDAKOOSTEFTNKLK-----EHAIVGLDNL- 133
 DB 1039 EELISKSSKDEYESQISLKEKLETRATYANDNVKNIKSELTKTRELAEALAAVNLK 1098
 QY 134 -----TDDNAORAI-----LKHANKDKG-----AALEKLFKAVENLS 167
 DB 1099 NELTKLETSEKALKVEKNEHLEKKEIKLEKATETTKQOLNSLRANLESLEKEHEDLA 1158
 QY 168 ---AAADDTLKNVAKELTSPYVHGNNSRKGCNASTNSADESVKPNLFEI-----SKKIT 219

DB 1159 AQLKTEEDIANKEROYNEIEIOLN---DEITSTQENESIKKRN-DELEGVAMKST 1213
 QY 220 ESNAAVLAVKEVELLASIDELATKAIGKIGNGLEANOSKNTSLSGAVASIDLAEK 279
 DB 1214 SEEOSNLKKESEIDALNLQIKELKK-----NETMEASLESISKVES----- 1255
 QY 280 LNVLKNEBKEKIDPAKOSTEFTNKLKSEHAIVGLDNLTDNNAORAILKHKAN----- 334
 DB 1256 -ETVKIKELQDCNFKREKVSLEDKLAKSE-----DKNSKYLELOKESKIKREL 1305
 QY 335 DKGAELEKLFKAVENLSKA-----AQDTLKNVAKEL 366
 DB 1306 DAKTELKIQLEKITNLKSAKESSELSRLKKTSEEKNAEBOL 1351
 RESULT 6
 NOFL_YEAST STANDARD; PRT; 944 AA.
 AC P32380;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NUPI PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110).
 GN NUPI OR SPC110 OR YDR356W OR D9476.3.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID:4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE=92176232; PubMed=1541631;
 RA Mirzayan C., Copeland C.S., Snyder M.;
 RT "The NUPI gene encodes an essential coiled-coil related protein that
 RT is a potential component of the yeast nucleoskeleton.";
 RL J. Cell Biol. 116:1319-1332(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94064779; PubMed=7503995;
 RA Klimartin J.V., Dyos S.L., Kershaw D., Finch J.T.;
 RT "A spacer protein in the Saccharomyces cerevisiae spindle poly body
 RT whose transcript is cell cycle-regulated.";
 RL J. Cell Biol. 123:1175-1184(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favellio A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucada T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifken L., Riles L., Tach A., Trevasz E., Vignati D.,
 RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
 CC ROLE IN CROSS-LINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT
 CC IS ESSENTIAL FOR GROWTH.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
 CC NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
 CC NUCLEOLUS.
 CC -1- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 DR EMBL: Z11582; CA477668.1; -
 DR EMBL: X73297; CA51733.1; -
 DR EMBL: U28372; AAB64791.1; -

DR PIR: S26710: S26710.
 DR PIR: S34288: S34288.
 DR SGD: S0002764: NUP1.
 KW Coiled coil: Nuclear protein; Phosphorylation.
 FT DOMAIN 164 791 COILED COIL.
 FT DOMAIN 54 59 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 726 731 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 742 747 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 731 944 ARG/TYR-RICH.
 SQ SEQUENCE 944 AA: 111781 MW: 04FAA074BB8A0BC8 CRC64;

Query Match 9.3%: Score 166.5; DB 1; Length 944;
 Best Local Similarity 23.4%; Pred. No. 0.55;
 Matches 100; Conservative 71; Mismatches 168; Indels 89; Gaps 22;

6 SGKDGASANSADSVKGNLTKTESNAV-VLAKEVEETLLASIDE-LATKAIGK 63
 91 SPRSGN-----VDSKRRNLIDDLKKDVPMSQPLKEQVEHQQKKRFRALSKILGK 145
 64 K-----IGNGLEANSKNTSLSGAVALSDIAEKLVLKN-ELKEKIDTAKQS 114
 146 RHITVANSISNKELYINEIKSLK-----HEIKELREKNDTNNYDTLEEDTDLKNRL 200
 115 TEFTNKLSEHAVLGDLNLTDDNAORAILKHKHKKGALELEKLFKAV-----ENLS 167
 201 QALEKELDAKNKIYV-SRAKVDHS--GCIEREROMEKRLAELEKLTVAQVLELENN 257
 168 -----KAAODTLKNAVKELETSPIVHGNSRRDGNAS-----TNSADE-SVKGN 210
 258 DVQSLKRSKEDEKLNLNMLENLE--LKSNAEKDQLEFKNELKRTNELNELKIKSDE 315
 211 L-TEISKITSENAVVLAKVEVELL-----ASIDELATKAIGKIGNGLEAN-OSKN 262
 316 MDQLKOKOKESRKLDELNELETKPESNGSSAKENELMKLNKTAE--LEEEISFTKN 373
 263 TSLSGAVALSDIAE-----KLVN-----LKNELKEKIDPAKOCSTFTNKLKSEH 310
 374 SOLIAKEGKLASLMAQITOLESKLNDORDSOLGSRPEELKKTNDKLOK-----DIRIR 428
 311 AVLGDLNLTDD-----DNAORAILKHKHANK-----DKGALELEKLFKAVENLSKAAODT 358
 429 TVSKDERIIDLOKKVKQLENDLFPYIKKTHSESKTITNELESKOKLIKILNDLKVAKQEK 488
 359 LKNAVKELE 366
 489 YSKMEKEL 496

RESULT 7

AC 092LT1; CAGA_HELPJ STANDARD; PRT; 1167 AA.

DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN)
 DE (CAG PATHOGENICITY ISLAND PROTEIN 26).
 GN CAGA OR CAI OR CAG26 OR JHP0495.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 NCBI_TaxID=85963;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir B.C., deJonge B.L., Carmel G.,
 RA Tummiano P.D., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RT Trist T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori."

RL Nature 397:176-180(1999).
 CC -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
 CC OR FUNCTION OF THE CYTOTOXIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE001483: AAD06073.1; -
 KW Antigen; Complete proteome.
 FT DOMAIN 246 249 POLY-THR.
 FT DOMAIN 882 889 POLY-ASN.
 SQ SEQUENCE 1167 AA: 129729 MW: FD5E8B81CEBD0F2 CRC64;

Query Match 9.1%: Score 163.5; DB 1; Length 1167;
 Best Local Similarity 26.2%; Pred. No. 0.95;
 Matches 102; Conservative 52; Mismatches 127; Indels 109; Gaps 20;

24 PNLTEIS-----KRTESNAVLAKEVEETLLASIDELATKAIGKIGNGLEA 72
 545 PNLNLAITSVVRDDEKLAKLSQPEANKLVKDFLSKKELVGKAL-----NPNKAV 599
 73 NQSKNTSLSGAV-----AISDLAEKLVNKELEKIDTAKOCSTFTNKLKSEHAV 127
 600 AEAKNT-----GNYDEVKQAKDL--EK--SLKKREBLEK-DVANKLESKGNKKMEAK- 649
 128 LGDLNLTDDNAQR-----AIIKKHANKDGALELEKLFKAVENLSKAAODTLKNAVKELETS 183
 650 -----SQANSQKDEIFALINKKANRDAARAI-----AAQNLKIGKRELSD 689
 184 PIVHGNSRRDGNASTNSADESVKGNLTKTESNAVLAKE-----VET 233
 690 KLENINNDLKD---FSKSPDEFKNGKN-KDPSKAEETLAKLGSVKDGIPEMYSKEN 745
 234 LLASIDELATKAIGKIGNGLEANSKNTSLSGAVALSDIAEKLVLKNKELEKID 293
 746 LNALNEF-----KNGKNKDFSKVTQ-----AKSDLENSIKVYIINRITDKVD 789
 294 TAKQ-----CSTFTNKLKSEHAVLGDLNLTDDN-AQRAILKHKHANKDGALELEKLFK 346
 790 NLNDAVSVAATGDFSG---VEQALADLKNFSKQLAQAQKNEDEFNGKNSA---LYQ 842
 347 AVEN-----LSKAAODTLKNAVKELE 366
 843 SVKNGVNGTIVNGLSKRAEATTLKSNPSDI 872

RESULT 8

AC 099323; MYSN_DROME STANDARD; PRT; 2017 AA.

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DE MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).
 GN MYO.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RX TISSUE=Embryo;
 RX MEDLINE=90349606; PubMed=2117279;
 RA Kechum A.S., Stewart C.T., Stewart M., Klehart D.P.;
 RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
 RT transcript: conserved sequences in the myosin tail and differential

```

RT splicing in the 5' untranslated sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: M5012; AAA28713.1; -.
PIR: A36014; A36014.
DR HSP: P08799; 1MD.
DR FLYBase: FBgn0005634; zfp.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ. 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ. 1.
DR SMART: SM00242; MISC; 1.
DR PROSITE: PS50096; IQ. 1.
KW Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW ATP-binding; Calmodulin-binding.
FT DOMAIN 1 829 MYOSIN HEAD-LIKE.
FT DOMAIN 886 859 IQ.
FT NP_BIND 225 232 COILED COIL (POTENTIAL).
FT DOMAIN 250 260 ATP.
FT DOMAIN 682 694 25 KDA/50 KDA JUNCTION.
FT DOMAIN 705 727 50 KDA/20 KDA JUNCTION.
FT DOMAIN 742 758 ACTIN-BINDING.
FT DOMAIN 1303 2017 REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT DOMAIN 1303 2017 LIGHT MEROMYOSIN (LMN).
FT DOMAIN 1971 2017 ALPHA-HELICAL TAILPIECE (LMN).
FT VAAPSLIC 1 45 GLOBULAR TAILPIECE.
SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;
-----
Query Match 9.0%; Score 162.5; DB 1; Length 2017;
Best Local Similarity 23.8%; Pred. NO. 1.9;
Matches 96; Conservative 63; Mismatches 139; Indels 101; Gaps 18;
-----
QY 26 LTEISKRKTESNAVY-----LAVEVE---TLTASIDELTKAIGKIG----- 66
Db 1108 VADLKEQENRRVQYDEMQAOLAKKEELLTQTLLRIDESKTKATAQAQELLESQAET 1167
QY 67 NNGLEANSKRTSLSGAYAISDLIAEKLVLKNEELKEKIDTAQCKSTFTNKLKSEHA 126
Db 1168 QEDLEAEKAAAK-----AEKYRDLSELLEALKNELLDLDTPTAAQ--QELRSKREGELA 1221
QY 127 VLGLDNLTDNAORAILKHHN-----KDKGAELKELFKAVENLSKA-----AQD 172
Db 1222 TL-----KSLSEETVNHGCVLADMRKHQSDELNSTINDOLENRKARKTVLEKAKG 1271
QY 173 TLKNAVELSPIYHGNNSRKGASTNSADESVKGNLTLEISKRTESNAVLAVK--E 230
Db 1272 TLEENADLATELRSSVNSROENDR-----RRKQAESQAELOVLAEL 1314
QY 231 VETLASIDELATAIGKIGN--NGLEANSKRTSLSGAYAISDLIAEKLVLKNEEL 288
Db 1315 IERARSELQEKCTK--LOOEAEENTINOLEEELKASAAVKSASNMESOLTEAOOLLE--EET 1372
= 3

```

```

QY 289 KEKIDTAQCKSTFTNKLKSEHAVIDGLNLTDMQORAI-----LKRHAKND 335
Db 1373 KQGLISRK-----LQOISEKEALQOLEEDEDKARNVERKLAETVTOMQIKKAEED 1427
QY 336 KG-AAELE-----KLFRVAVENLSK-----AADDTLKNAVELTS 368
Db 1428 ADLAKLEEGKKRLKNDIDALERQYKELIAQNDRLDKSRKKIQS 1471
-----
RESULT 9
MPL YEAST STANDARD; PRT: 1875 AA.
AC 002455;
ID 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN-LIKE PROTEIN MLP1.
GN MLP1 OR YKR095W OR YKRA15.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C.
RC MEDLINE=93247549; PubMed=8483450;
RA Koelling R., Nguyen T., Chen E.Y., Botstein D.;
RT "A new yeast gene with a myosin-like heptad repeat structure."
RL Mol. Genet. 237:359-369(1993).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=94205265; PubMed=8154186;
RA Bou G., Bsteden P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT "The complete sequence of a 15,820 bp segment of Saccharomyces
RT cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three
RL new open reading frames."
RL Yeast 9:1349-1354(1993).
CC -1- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
CC REPAIR.
CC -1- SIMILARITY: SOME, TO THE TPR ONCOGENE.
CC -1- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL1".
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: L01992; AAA3783.1; -.
DR EMBL: X73541; CA51948.1; -.
DR EMBL: Z28320; CAAB2174.1; -.
DR PIR: S38173; S38173.
DR SGD: S0001803; MLP1.
KW Coiled coil; DNA repair.
FT DOMAIN 69 487 COILED COIL (POTENTIAL).
FT DOMAIN 531 1678 COILED COIL (POTENTIAL).
FT DOMAIN 1834 1866 COILED COIL (POTENTIAL).
FT CONFLICT 301 301 R -> A (IN REF. 1).
SQ SEQUENCE 1875 AA; 218455 MW; 683AD034C9066867 CRC64;
-----
Query Match 8.9%; Score 159.5; DB 1; Length 1875;
Best Local Similarity 21.3%; Pred. NO. 2.4;
Matches 96; Conservative 80; Mismatches 155; Indels 119; Gaps 18;
-----
QY 26 LTEISKRKTESNAVYLAVEVETLASIDELATAIGKIGNNGLEANSKRTSLSGAY 85
Db 647 ISQITRESTEN--MSLLNNEIDLYDSKDISIK-LGKRSKRILAEERK---LLSNLT 700

```

[illegible]

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M88098; AAA29744.1; -
KW Malaria; Receptor; Membrane.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CNC64;

Query Match 8.88; Score 157.5; DB 1; Length 1251;
Best Local Similarity 20.08; Pred. No. 1.9;
Matches 106; Conservative 77; Mismatches 151; Indels 181; Gaps 21;

QY 9 DGNASANSADSESVK-----GPNLEFISKITESSNAVYLA-----KEVEFLASIDELAT 58
DB 671 EGHGNVQVJENIKELVDENNLSIDLIKQATGKEEIOKTHSTLKNKAKTILGHVDTSA- 729
DB 59 KAIGKKT-----GNNGLEA-NOSKNTSLISGAYAISDLI 91
DB 730 KYVGKITPPLALVELLDGAKLKTAEKLPESKNNVYLETENNSKNTNEDVHKNQDAY 789
QY 92 AEKLVNKNELKEKIDTAKOCST-----EFTNKLKS----- 123
DB 790 KVALEILAH---SDEIDTKOKSSKLIEMGNQIYKVLINQKKNISSIKSEEAAYSVK 846
QY 124 -----EHAVLG-----LDNLTDDNACRAILKKRANKDKGAEELE 157
DB 847 IGVNYSKKHSELKITSCKSDSYDNIILEKOTELQNLRSNFTQE---KTMND---SKLE 900
QY 158 KLFKAVENTLSKAODPLKNAVKELTSPYHGNNSRKDGNAISNADSESKGPNLEISKK 217
DB 901 KI-----KIDFESLKNALKTLEGE-VNALKASSDNEHVQSKSEPV-NPALSEIEKE 950
QY 218 ITESNAVLAKEV---ETLLASIDELATKAIGKIGN-----NGLEAN----- 258
DB 951 EKDIDSLNALDELKKGTCESRYKLLKDVYKESIDDTLINTIENNVAYIAYIKK 1010
QY 259 -----OSKNTSLISGAYAISDLIAEKL-----NV 282
DB 1011 NEDTVQVQVLTNEHENTKOVNSHEPTNEDKNSSEELTKAVTSKTIISKLGVITTEV 1070
QY 283 LKNEELKEKIDTAKOCSTFTNKLKSEHAVLGLDNLTDNAQRALLKHKANKAAEELE 342
DB 1071 NENTEMNTIESSAKTEEALY-NELKNKTSLSLEIYQTSNEVYKLEMKSNADK---YIDVS 1126
DB 343 KLFKAVENTLSKA-----AODTLKNAYKEL 366
DB 1127 KIFNTVLDTKQNSIVTNGHSINNVAKDKLKGKLQEL 1161

RESULT 12
LNG1_DROME STANDARD; PRT; 1639 AA.
AC P15215; Q24373; Q9VT18;
DT 01-APR-1990 (Rel. 14, Created).
DT 01-NOV-1997 (Rel. 35, Last sequence update).
DT 20-AUG-2001 (Rel. 40, Last annotation update).
DE LAMININ GAMMA-1 CHAIN PRECURSOR (LAMININ B2 CHAIN).
GN LAMB2 OR LAMC1 OR LAMG1 OR CG3322.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S; AND OREGON-R;
RX MEDLINE=91299161; PubMed=1840513;
RA Chi H.-C., Jumianga D., Wang S.-Y., Hui C.-F.;
RT "Structure of the Drosophila gene for the laminin B2 chain.";
RL DNA Cell Biol. 10:451-466(1991).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=89109164; PubMed=2912972;
RA Chi H.-C., Hui C.-F.;
RT "Primary structure of the Drosophila laminin B2 chain and comparison with human, mouse, and Drosophila laminin B1 and B2 chains.";
RL J. Biol. Chem. 264:1543-1550(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90037237; PubMed=2808533;
RA Montell D.J., Goodman C.S.;
RT "Drosophila laminin: sequence of B2 subunit and expression of all three subunits during embryogenesis.";
RL J. Cell Biol. 109:2441-2453(1989).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Burton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P., Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iredregam C., Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshire A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [5]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [6]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [7]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [8]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [9]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [10]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [11]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [12]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [13]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [14]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [15]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [16]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [17]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [18]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [19]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [20]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [21]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [22]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [23]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [24]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [25]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [26]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [27]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [28]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [29]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [30]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [31]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [32]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [33]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [34]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [35]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [36]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [37]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [38]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [39]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [40]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [41]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [42]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [43]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [44]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [45]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [46]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [47]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [48]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [49]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [50]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [51]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [52]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [53]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [54]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [55]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [56]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [57]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [58]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [59]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [60]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [61]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [62]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [63]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [64]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [65]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [66]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [67]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [68]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [69]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [70]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [71]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [72]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [73]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [74]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [75]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [76]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [77]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [78]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [79]
RP SEQUENCE OF 344-1639 FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=88303364; PubMed=3405777;
RA Chi H.-C., Hui C.-F.;
RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL Nucleic Acids Res. 16:7205-7205(1988).
RN [80]

[illegible]

Query Match	Best Local Similarity	8.7%;	Score 157;	DB 1;	Length 1639;
Matches 104;	Conservative	61;	Mismatches 158;	Indels 108;	Gaps 21;
2	ACNNGSKGNSANSAD-----	ESVKGPNIETSKITESNAVYLAVKEVETLLASIDE	55		
Db	1107 AADNSGGGQYAEYVIDLHKHLDVSR-EHLVSADKFOADANGETIDRARKONTIIDQITE	1165			
Qy	56 LATKAIGAKI-----GNNGLEANOSKNTSLSGAYAISDL-----IAEKLNV-----L	98			
Db	1166 NAKKTELQDALDLINDEGAQALARAKEKSVFEGQOSEQSLDISREARALADKLESAQFDL	1225			
Qy	99 KN-EELKEKIDTAKOCS-----TEFTNKKLSHNAVJGLDNLITDNDNAORAILKHKANKD	150			
Db	1226 KNAKDAKDAVEKAHOLAKSAIDLOLKISTELTSE---VGLT-----LSHVKOS	1270			
Qy	151 KGA---AELEKFKAVENLSKRAOQTILNAVVELSPIYHGNSKRKDGNSATNSADESVK	207			
Db	1271 LGTVVOTSKEARAKNEVEYDTRL--TLINDNARQOPPEIDISQDKDVAANERADELKL	1328			
Qy	208 GPNIETSKITESNAVYLAVKEVETLLASIDELTKAIGKKGNNNGLEANOSKNTSLSL	267			
Db	1329 --QITELS-----NSNGELFADEFTQEL-----TALLKR-----AEQOGLDIELLE	1370			
Qy	268 GAYAISDLIAEKL-----NVLK-----NEELKEKIDTAKOCSFTFTN	304			
Db	1371 RAKAHDATKATVAEBOGDWTLKEANTTYEKLAFQSGDVORSSASAKALQTPNIEKEION	1430			
Qy	305 K---LKESENAVJGLD-----NLTDNNAORAILK--HNAKKDGAALKEKLFKAVENLSKAA	355			
Db	1431 AESLISQAEALDGNKANKNAEKKNAEOELKYTAQASKD-----AELIRKKAETKYVA	1484			
Qy	356 QDTLKNAYKEL	366			
Db	1485 ARNLFEADQL	1495			

A60D_DROME STANDARD; PRT; 1013 AA.
 ID A60D_DROME
 AC P91927; G9W160;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CALCIUM-BINDING MITOCHONDRIAL PROTEIN ANON-60DA.
 GN ANON-60DA OR CG4589.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Ephygryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephygryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC NBI_Taxid-7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gccayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fliscmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Harris N.L., Harvey D., Helman T.J., Herndon J.R., Houck J.,
 RA Hostin D., Houston K.A., Heiman T.J., Herndon J.R., Houck J.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [12]
 RP SEQUENCE OF 626-944 FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=99168769; PubMed=10071211;
 RA Caggese C., Regone G., Perrini B., Moschetti R., de Pinto V.,
 RA Calzi R., Barsanti P.;
 RT "Identification of nuclear genes encoding mitochondrial proteins:
 RT isolation of a collection of D. melanogaster cDNAs homologous to
 RT sequences in the Human Gene Index database.";
 RL Mol. Gen. Genet. 261:64-70(1999).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS IN POSITIONS 920 AND 930.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use.

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE003464; AAP47217.1; -
 DR EMBL: Y10912; CA71853.1; ALT_FRAME.
 DR FlyBase: FBgn0019886; CG4598.
 DR InterPro: IPR001813; 60S_ribosomal.
 DR InterPro: IPR002024; Bacterioferitin.
 DR InterPro: IPR003871; DUF223.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000861; REM_repeat.
 DR InterPro: IPR002555; RFL.
 DR InterPro: IPR001236; 1dh.
 DR Pfam: PF00428; 60S_ribosomal; 1.
 DR Pfam: PF01334; Bacteriofer; 1.
 DR Pfam: PF02721; DUF223; 1.
 DR Pfam: PF00036; efhand; 2.
 DR Pfam: PF02185; HRI; 1.
 DR Pfam: PF00056; 1dh; 1.
 DR Pfam: PF01605; RFL; 1.
 DR SMART: SM00054; EFh; 2.
 DR PROSITE: PS00018; EF_HAND; 2.
 KW Mitochondrion; Calcium-binding; Repeat.
 FT CA_BIND 700 711 EF_HAND 1 (POTENTIAL).
 FT CA_BIND 783 794 EF_HAND 2 (POTENTIAL).
 FT CONFLICT 687 687 E -> Q (IN REF. 2).
 FT CONFLICT 690 690 K -> G (IN REF. 2).
 FT CONFLICT 717 717 Q -> H (IN REF. 2).
 FT CONFLICT 736 736 E -> D (IN REF. 2).
 FT CONFLICT 740 740 K -> R (IN REF. 2).
 FT CONFLICT 761 761 K -> E (IN REF. 2).
 FT CONFLICT 943 943 G -> S (IN REF. 2).
 SQ SEQUENCE 1013 AA; 113579 MW; 8C710380263F262E CRC64;
 Query Match 8.7%; Score 156.5; DB 1; Length 1013;
 Best Local Similarity 22.6%; Pred. No. 1.7;
 Matches 81; Conservative 62; Mismatches 141; Indels 75; Gaps 12;
 QY 29 ISKTTESAVVLAKEVEETLLASIDELATYKAIGKNGNGLANOSKNTSLSGVAIS 88
 DB VKEPREERAAKLLVNRNKMISQIDNVL-----NDLEAHOIKOAESSDYAAS 658
 QY 89 DLAELKLVNKEELKEKIDPAKOCSTFTNKIKSEHAVLGIDNITDNNAGRAILKKHAN 148
 DB SPVPEPQWVHIDELVATIRKKEASDERFRVGD-----LVKKDAD 702
 QY 149 KDKGAELKELEKFAVENLSKAAODTLKNAVELT-----SPVHGNSRKDGN 196
 DB KD-GVIVNEITKAVQSIDREATNIDKKQLEPELKLKSLARRRHEIVAIIDLMNNIK 761
 QY 197 ASTNADSVGPNITELSKITNESNAVVAKEVEETLLASIDELATYKAIGKNGNGL 256
 DB VLKETSDEA-RLEKHEALEKFDADKDGVTVDNIRKLVES-----IGRDNK 808
 QY 257 ANOSKNTSLSGAAYASIDLA--EKLNVKNEELKEKIDTKOCSTFTNKIKSE--HAV 312
 DB 809 LSDK-----AIEELISLDKEOVLAE--QAIERAIKASKEAEKLSSEVDKAD 855
 QY 313 LGDNLITDNNQRA-----ILKHKANKDGAELKELEKFAVENLSKAAODTLKNAVEL 366
 DB 856 KDLKLVNDIHDISAKETIODIANEMRDKREIVPDKAKELKA-EPAFKDTATLTQMDNADL 913
 RESULT 14
 ID MS2_SCHPO STANDARD; PRT; 1526 AA.
 AC G9US16; P78969;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE MYOSIN TYPE II HEAVY CHAIN 1.
 GN MYO2 OR SPCC645.05C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC STRAIN-972;
 RX MEDLINE=98075862; PubMed=9415380;
 RA May K.M., Watts F.Z., Jones N., Hyams J.S.,
 RT "Type II myosin involved in cytokinesis in the fission yeast,
 RL Schizosaccharomyces pombe."
 RN Cell Motil. Cytoskeleton 38:385-396(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA Wood V., Rajandream M.A., Barrell B.G., Rieger M.,
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR CELL DIVISION. IT IS A COMPONENT OF THE
 CDC12 'SPOT', A STRUCTURE THOUGHT TO MARK THE SITE OF SEPTATION.
 CC MAY WORK IN CONJUNCTION WITH MYO3.
 CC -1- SUBUNIT: BINDS TO CDC4 AND RLC1.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U75357; AAC49908.1; -
 DR EMBL: AL048498; CAB39901.1; -
 DR HSSP: P08799; 1MD.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ.1.
 DR Pfam: PF00063; myosin_head.1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head.1.
 DR SMART: SM00015; IQ.1.
 DR SMART: SM00242; MYSC.1.
 DR PROSITE: PS00096; IQ.1.
 DR Myosin: Actin-binding; ATP-binding; coiled coil; Calmodulin-binding;
 KW Alkylation.
 KM
 FT DOMAIN 1 757 MYOSIN HEAD-LIKE.
 FT DOMAIN 758 787 IQ.
 FT NP_BIND 875 1244 COILED COIL (POTENTIAL).
 FT FT_BIND 170 177 ATP (POTENTIAL).
 FT DOMAIN 634 656 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 734 748 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 674 674 ALKYLATION (BY SIMILARITY).
 FT CONFLICT 1337 1337 S -> R (IN REF. 1).
 SQ SEQUENCE 1526 AA; 176430 MW; D71D51D6578192BA CRC64;

Query Match 8.6%; Score 154; DB 1; Length 1526;
 Best Local Similarity 23.3%; Pred. No. 3.3; Indels 74; Gaps 18;
 Matches 88; Conservative 76; Mismatches 139;

OY 8 KDSGANSANDESVPKPN-LTEISKRTTESNAVLAVKEV-----ETLLASIDE----- 55
 DB 842 KQNSKSEVERDLVETNNSTLAVENLLTERATLADKEELIKRTQERLANIEDSFSTKQ 901
 OY 56 ----LATAIGKRTGNNGLEANOSKNTSLSGAVYASDLIAEKLAVLKNELEKIDYAK 111
 DB 902 QNENLQRESASLKQINNELESELEKTS-----KVETLLSEQ-----NELKEKLSLEE 949
 OY 112 QGSTFTFNKLS--EHAIVGLDNLDTDMNQ-RAILKKHAKKDGAELEKLFKAVENLSK 168

DB 950 KDLDTKGELSLRENNATYVLSAEKAFNEOCKSLQETITVTD---AELDKLTITSDYKT 1006
 OY 169 AADDTLNAKVELTSPYVHGNNSKRDGNASTNSADESVKGNELISKRTTESNAVLA- 227
 DB 1007 EIQE-----MRLTNQMNKESITQDEGSL-----ESLKR-----VKLLERESTLSD 1049
 OY 228 -----VKVEETLLASIDELATAIGKRTGNNGLEANOSKNTSLSGAVYASDL-IAEKL 280
 DB 1050 VSLKQKKEELSVLKGVOELTINLNEKV--NYLEADVOKLPKIKKELESINDQDLYQL 1107
 OY 281 NVLKNELKEKIDTAKCCSTFTNKLKS-----EHAIVGLDNLDTDMAQRIILKK-IANK 334
 DB 1108 QATRNKELEAKV--KCC---LNNIKSLTKELENKEEKCQNLSDASLKYIEIQEIHENL 1160
 OY 335 DKGAELKELFKAVENL 351
 DB 1161 LKAVSDLENKTKKEGL 1177

RESULT 15
 CEAK_ECOLI
 ID CEAK_ECOLI STANDARD; PRT; 548 AA.
 AC 047502; P75615;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE COLICIN K.
 GN CKA.
 OS Escherichia coli.
 OG Plasmid ColK-K235.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K49;
 RX MEDLINE=96074330; PubMed=7592493;
 RA Pils H., Braun V.,
 RT "Strong function-related homology between the pore-forming colicins K
 RT and 5."
 RT J. Bacteriol. 177:6973-6977(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Izard J., Chartier M., Baly D.,
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF
 CC TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING
 CC TO DISSIPATION OF CELLULAR ENERGY.
 CC -1- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
 CC AGAINST ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
 CC -1- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X87834; CAA61099.1; -
 DR EMBL: U27452; AAB41288.1; -
 DR InterPro: IPR000293; Channel_colicin.
 DR Pfam: PF01024; Colicin.1.
 DR PRINTS: PR00280; CHANCOLICIN.
 DR PRODOM: PD002657; CHANCOLICIN.
 DR PROSITE: PS00276; CHANNELCOLICIN.1.
 KW Antibiotic; Bacteriocin; Plasmid; Transmembrane.
 FT TRANSMEM 505 525 POTENTIAL.
 FT CONFLICT 155 155 S -> G (IN REF. 2).
 SQ SEQUENCE 548 AA; 59661 MW; 2E6768900808CFF6 CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:54 ; Search time 124.19 Seconds
(without alignments)
433.434 Million cell updates/sec

Title: US-09-596-746A-38
Perfect score: 1797
Sequence: 1 MACNNSCKDGNASANSADSES.....ENLSKAADTLKNAVKELTS 368

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.invertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description
1	901	50.1	192	2	Q9RS52	Q9RS52 borrelia bu
2	901	50.1	209	2	Q44717	Q44717 borrelia bu
3	899	50.0	191	2	Q9S3P1	Q9S3P1 borrelia bu
4	889	49.5	192	2	O31117	O31117 borrelia bu
5	663	36.9	194	2	O31122	O31122 borrelia bu
6	661	36.8	193	2	P94234	P94234 borrelia bu
7	653.5	36.4	212	2	Q44705	Q44705 borrelia bu
8	646	35.9	202	2	O08231	O08231 borrelia bu
9	646	35.9	211	2	Q44720	Q44720 borrelia bu
10	644	35.8	201	2	Q9R0R8	Q9R0R8 borrelia bu
11	642	35.7	193	2	P94237	P94237 borrelia bu
12	639	35.6	191	2	Q44726	Q44726 borrelia bu
13	623.5	34.7	192	2	Q9R7B1	Q9R7B1 borrelia bu
14	623.5	34.7	193	2	Q9R8S3	Q9R8S3 borrelia bu
15	623	34.7	185	2	O31123	O31123 borrelia bu
16	621.5	34.6	192	2	Q9S3P3	Q9S3P3 borrelia bu
17	620	34.5	200	2	Q9XDH4	Q9XDH4 borrelia bu
18	619	34.4	193	2	P94236	P94236 borrelia bu
19	616	34.3	209	2	Q44883	Q44883 borrelia bu

20	614	34.2	191	2	Q9S3P0	Q9S3P0 borrelia bu
21	612	34.1	188	2	Q9XDH3	Q9XDH3 borrelia bu
22	610.5	34.0	203	2	O50620	O50620 borrelia af
23	609.5	33.9	200	2	Q9R0R9	Q9R0R9 borrelia bu
24	609.5	33.9	205	2	P96505	P96505 borrelia af
25	607.5	33.8	193	2	O31115	O31115 borrelia bu
26	607.5	33.8	201	2	P96571	P96571 borrelia ja
27	607.5	33.8	212	2	Q44670	Q44670 borrelia af
28	605.5	33.7	191	2	P70818	P70818 borrelia bu
29	605.5	33.7	194	2	P94229	P94229 borrelia bu
30	605.5	33.7	201	2	P96573	P96573 borrelia ja
31	605	33.7	178	2	O44979	O44979 borrelia bu
32	605	33.7	178	2	O44995	O44995 borrelia bu
33	603.5	33.6	210	2	Q44719	Q44719 borrelia bu
34	602.5	33.5	201	2	P96516	P96516 borrelia ja
35	601.5	33.5	192	2	Q9S3P2	Q9S3P2 borrelia bu
36	601.5	33.5	201	2	P96572	P96572 borrelia ja
37	600	33.4	190	2	P70819	P70819 borrelia bu
38	599	33.3	184	2	Q9S504	Q9S504 borrelia bu
39	598.5	33.3	191	2	O31120	O31120 borrelia bu
40	596.5	33.2	182	2	Q9R7B2	Q9R7B2 borrelia bu
41	596.5	33.2	212	2	Q44718	Q44718 borrelia bu
42	594.5	33.1	201	2	P96514	P96514 borrelia ja
43	594	33.1	181	2	O34124	O34124 borrelia bu
44	592	32.9	184	2	O34120	O34120 borrelia bu
45	591	32.9	211	2	Q9Z6C7	Q9Z6C7 borrelia va

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	192 AA.
ID	Q9RS52	Q9RS52		
AC	Q9RS52			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	OUTER SURFACE PROTEIN C (FRAGMENT).			
GN	OSPC.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID=139;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=OC5;			
RX	MEDLINE=99091544; PubMed=9872945;			
RA	Wang I.N., Dykhuizen D.E., Olu W., Dunn J.J., Bosler E.M., Luft B.J.;			
RT	"Genetic diversity of ospC in a local population of Borrelia burgdorferi sensu stricto."			
RL	Genetics 151:15-30(1999).			
DR	EMBL: AF029864; AAB86547.1; "			
DR	InterPro: IPR001800; Lipoprotein_6.			
DR	Pfam: PF01441; Lipoprotein_6; 1.			
FT	ProDom: PD001149; Lipoprotein_6; 1.			
FT	NON_TER 1			
FT	NON_TER 192			
SQ	SEQUENCE 192 AA; 20555 MW; FFL6D409D58C01C5 CRC64;			
Query Match	50.1%;	Score 901;	DB 2;	Length 192;
Best Local Similarity	98.9%;	Pred. No. 3;	6e-35;	
Matches 144;	Conservative	2;	Mismatches 0;	Indels 0;
			Gaps 0;	
QY	1	MACNNSCKDGNASANSADSESVKGPNTLEISKITTESNAVYLAWEVETLLASIDELATKA	60	
DB	7	ISCNNSCKDGNASANSADSESVKGPNTLEISKITTESNAVYLAWEVETLLASIDELATKA	66	
QY	61	IGKKGNGLEANSKNTSLSGAYAISDLIAEKLNVLNKEELKEKIDTAACOSTEFTNK	120	
DB	67	IGKKGNGLEANSKNTSLSGAYAISDLIAEKLNVLNKEELKEKIDTAACOSTEFTNK	126	
QY	121	LKSEHAVLGIDNLTDDNAQAIIKKHANKDKGAEELEKFRAYENLSKAADTLKNAVKE	180	

```

Db 127 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 186
QY 181 LTSPITV 186
Db 187 LTSPITV 192

RESULT 2
ID 044717 PRELIMINARY: PRT: 209 AA.
AC 044717:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE OSpC.
GN OSpC.
OR Borrelia burgdorferi (Lyme disease spirochete).
OR Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OR NCBI_Taxid=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RX MEDLINE=94314484; PubMed=8039931;
RA Stevenson B., Bockenstedt L.K., Barthold S.W.;
RT "Expression and gene sequence of outer surface protein C of Borrelia
RT burgdorferi reisolated from chronically infected mice.";
RL Infect. Immun. 62:3568-3571(1994).
DR EMBL: U04240; AAC45538.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
SQ SEQUENCE 209 AA; 22393 MW; 3707A47DAA736FCA CRC64;

Query Match 50.1%; Score 901; DB 2; Length 209;
Best Local Similarity 98.9%; Pred. No. 4e-35;
Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNSGKDGNASANSADSVKGPMLTEISKRTTESNAVVLAVKEVEETLLASIDELATKA 60
DB 17 ISCNNSKDGNASANSADSVKGPMLTEISKRTTESNAVVLAVKEVEETLLASIDELATKA 76
QY 61 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLVNKNELEKIDPAKOCSTEEFTNK 120
DB 77 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLVNKNELEKIDPAKOCSTEEFTNK 136
QY 121 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180
DB 137 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 196
QY 181 LTSPITV 186
DB 197 LTSPITV 202

RESULT 3
ID 09S3P1 PRELIMINARY: PRT: 191 AA.
AC 09S3P1:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OR Borrelia burgdorferi (Lyme disease spirochete).
OR Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OR NCBI_Taxid=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26691;
RX MEDLINE=96296448; PubMed=8709845;
RA Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT "Evidence for lateral transfer and recombination in OSpC variation in

```

```

RT Lyme disease Borrelia.";
RL Mol. Microbiol. 18:257-269(1995).
DR EMBL: LA2894; AAB37002.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 191
SQ SEQUENCE 191 AA; 20340 MW; 8CA1A64CF17AEDBF CRC64;

Query Match 50.0%; Score 899; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 4e-35;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CNNSGKDGNASANSADSVKGPMLTEISKRTTESNAVVLAVKEVEETLLASIDELATKA 62
DB 1 CNNSGKDGNASANSADSVKGPMLTEISKRTTESNAVVLAVKEVEETLLASIDELATKA 60
QY 63 KIGNGNGLEANSKNTSLSGAVAIISDLIAEKLVNKNELEKIDPAKOCSTEEFTNK 122
DB 61 KIGNGNGLEANSKNTSLSGAVAIISDLIAEKLVNKNELEKIDPAKOCSTEEFTNK 120
QY 123 SEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 182
DB 121 SEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180
QY 183 SPITV 186
DB 181 SPITV 184

```

```

RESULT 4
ID 031117 PRELIMINARY: PRT: 192 AA.
AC 031117:
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSpC.
OR Borrelia burgdorferi (Lyme disease spirochete).
OR Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OR NCBI_Taxid=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OC7;
RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF029866; AAB86549.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 20684 MW; 1F0C6BB6E291F6B CRC64;

Query Match 49.5%; Score 889; DB 2; Length 192;
Best Local Similarity 97.8%; Pred. No. 1.3e-34;
Matches 182; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNSGKDGNASANSADSVKGPMLTEISKRTTESNAVVLAVKEVEETLLASIDELATKA 60
DB 7 ISCNNSKDGNASANSADSVKGPMLTEISKRTTESNAVVLAVKEVEETLLASIDELATKA 66
QY 61 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLVNKNELEKIDPAKOCSTEEFTNK 120
DB 67 ICKKIGNNGLEANSKNTSLSGAVAIISDLIAEKLVNKNELEKIDPAKOCSTEEFTNK 126
QY 121 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 180
DB 127 LKSEHAVLGDLNLTDDNQRALILKHKANKDKGAELKLFKAVENLSKAAODTLKNAVKE 186

```


OY 178 VKELTSPVHGN 189
 DB 197 VKELTSPVAEN 208

RESULT 8

ID 008231 PRELIMINARY; PRT; 202 AA.
 AC 008231;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia tanukii.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=56146;
 RN [1]
 SEQUENCE FROM N.A.
 RM STRAIN=FI81T;
 RM Masuzawa T., Komikado T., Fukui T., Yanagihara Y.;
 RT "B. tanukii (Strain FI81t) ospC, partial cds."
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB000354; BAI19087.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 202
 SQ SEQUENCE 202 AA; 21536 MW; F309F684C68961E7 CRC64;

Query Match 35.9%; Score 646; DB 2; Length 202;
 Best Local Similarity 70.7%; Pred. No. 2.2e-23;

Matches 133; Conservative 22; Mismatches 31; Indels 2; Gaps 2;

OY 1 MACNNSGKRG-NASANSADSVKGNLTETISKITESNNAVYLAKEVEETLLASIDELATK 59
 DB 11 ISCNNSGKGDGASINPVDSEKAGPULTEISKITDSNAIVLAKEVEETLLSIDELANK 70
 OY 60 AIGKRIINGNGLEANOSKNTSLSGAVASIDLAETKLVNKN-EELKEKIDTAQOCSTFT 118
 DB 71 AIGKRIIONNGDLTLDKNAKSLSGAVYSTLTETKLNGLKNSDLEKEIEKKKCSAFT 130
 OY 119 NKLKSEHAVLGIDNLTDDNAORAILKHKANKDKGAELKFKAVENLSKAQDTLKNAV 178
 DB 131 NKLKSHOVLGEAATDDAKKAILKHTYTKDKGAEEFKLFSVEILIKAAQELANSI 190
 OY 179 KELTSPV 186
 DB 191 KELTSPV 198

RESULT 9

ID 044720 PRELIMINARY; PRT; 211 AA.
 AC 044720;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OSPC.
 GN OSPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=25015;
 RC MEDLINE=95154673; PubMed=7851744;
 RA Stevenson B., Barthold S.W.;
 RT "Expression and sequence of outer surface protein C among North
 American isolates of Borrelia burgdorferi."
 RL FEMS Microbiol. Lett. 124:367-372(1994).
 DR EMBL: U04282; AAC45540.1; -

DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 211 AA; 22603 MW; 63984B6D8743ED5 CRC64;

Query Match 35.9%; Score 646; DB 2; Length 211;
 Best Local Similarity 72.3%; Pred. No. 2.3e-23;
 Matches 136; Conservative 18; Mismatches 32; Indels 2; Gaps 2;

OY 1 MACNNSGKDG-NASANSADSVKGNLTETISKITESNNAVYLAKEVEETLLASIDELATK 59
 DB 17 ISCNNSGKGDGASINPVDSEKAGPULTEISKITDSNAIVLAKEVEETLLSIDELATK 76
 OY 60 AIGKRI-GNNGLEANOSKNTSLSGAVASIDLAETKLVNKN-EELKEKIDTAQOCSTFT 118
 DB 77 AIGKRIHONNGDLENNHNGSLAGAVASITLTQKLGKLNELKEKIAVKKCSSEFT 136
 OY 119 NKLKSEHAVLGIDNLTDDNAORAILKHKANKDKGAELKFKAVENLSKAQDTLKNAV 178
 DB 137 NKLKSHTELGKQADQADAKKAILKHTYTKDKGAEEIDKLFKAVENLSKAQELNSV 196
 OY 179 KELTSPV 186
 DB 197 KELTSPV 204

RESULT 10

ID 09RQ8 PRELIMINARY; PRT; 201 AA.
 AC 09RQ8;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
 GN OSPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=3B6';
 RC MEDLINE=20002545; PubMed=10531219;
 RA Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
 RT "Population dynamics of a naturally occurring heterogeneous mixture of
 Borrelia burgdorferi clones."
 RL Infect. Immun. 67:5708-5716(1999).
 DR EMBL: AF074465; AAD23912.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 201
 SQ SEQUENCE 201 AA; 21457 MW; 806F198295101B07 CRC64;

Query Match 35.8%; Score 644; DB 2; Length 201;
 Best Local Similarity 73.4%; Pred. No. 2.7e-23;
 Matches 135; Conservative 22; Mismatches 25; Indels 2; Gaps 2;

OY 1 MACNNSGKDG-NASANSADSVKGNLTETISKITESNNAVYLAKEVEETLLASIDELATK 60
 DB 17 ISCNNSGKGDGASINPVDSEKAGPULTEISKITDSNAIVLAKEVEETLLSIDELATK 76
 OY 61 IGKRI-GNNGLEANOSKNTSLSGAVASIDLAETKLVNKN-EELKEKIDTAQOCSTFT 118
 DB 77 IGKRIIONNGDLAEGHNGTLLAGAYTISKILITQKIDGLKNSKLEKIEKAKKCSDEPT 136
 OY 119 NKLKSEHAVLGIDNLTDDNAORAILKHKANKDKGAELKFKAVENLSKAQDTLKNAV 178
 DB 137 KKLGEHAOLGIEIVTDENAKKAILITTDKADKGAAELKFKAVENLSKAQELANSV 196

Oy 179 Kelt 182
 DB 197 Kelt 200

RESULT 11
 ID P94237

PRELIMINARY: PRT: 193 AA.

AC P94237: 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
 GN OUTR SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxId-139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-25015;
 PX MEDLINE-96296448; PubMed-8709845;
 RT Livy I., Gibbs C.P., Schuster R., Dorner F.;
 "Evidence for lateral transfer and recombination in OspC variation in
 Lyme disease Borrelia."
 RL Mol. Microbiol. 18:257-269(1995).
 DR EMBL: L42898; AAB37007.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT SEQUENCE 193 AA: 20677 MW: C9500959E13590D CRC64;

Query Match 35.7%; Score 642; DB 2; Length 193;
 Best Local Similarity 72.6%; Pred. No. 3.2e-23;
 Matches 135; Conservative 16; Mismatches 33; Indels 2; Gaps 2;

Oy 3 CUNSGKDG-N-ASANSADSVKGNLTETISKITESNAVLAKEVEETLLASIDELATKAI 61
 DB 1 CUNSGKDGNAASNPADSVKGNLTETISKITESNAVLAKEVEETLLASIDELATKAI 60
 Oy 62 GKRI-GNNGLEANQKNTSILSGAVAISDLIAEKLNVKNEELKEKIDTAOKCSTFTNK 120
 DB 61 GKRIHONNGIDTENNHNGLAGAVASTLTITKGLKNEELKEKIAAVKCSSEFTNK 120
 Oy 121 LKSEHAVLGIDNLTDDNAORAILKHKANKGAAELEKLFKAVENTLSKAQDTLKNVKE 180
 DB 121 LKSSHRELCKQADODDADAKKAILRTINTKDKGAELEKLFKPYENLSKAKEMLSNVKE 180
 181 LTSPIV 186
 DB 181 LTSPV 186

RESULT 12
 ID Q44726 PRELIMINARY: PRT: 191 AA.
 AC Q44726: 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
 GN OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxId-139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-297;
 RA MEDLINE-94314437; PubMed-8039891;
 RA Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;
 "Humoral immune response to outer surface protein C of Borrelia
 burgdorferi in Lyme disease: role of the immunoglobulin M response in

RT the serodiagnosis of early infection.";

RL Infect. Immun. 62:3213-3221(1994).

DR EMBL: U08284; AAA21460.1;-

DR InterPro: IPR001800; Lipoprotein_6.

DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT SEQUENCE 191 AA: 20150 MW: C49A4030F0A28717 CRC64;

Query Match 35.6%; Score 639; DB 2; Length 191;
 Best Local Similarity 73.9%; Pred. No. 4.4e-23;
 Matches 136; Conservative 20; Mismatches 26; Indels 2; Gaps 2;

Oy 5 NSGKDGNAASNSADSVKGNLTETISKITESNAVLAKEVEETLLASIDELATKAI 64
 DB 1 NSGKDGNAASNSADSVKGNLTETISKITESNAVLAKEVEETLLASIDELATKAI 60
 Oy 65 I-GNNGLEANQKNTSILSGAVAISDLIAEKLNVKNEELKEKIDTAOKCSTFTNK 122
 DB 61 IONGGLAVAGHNGTLLAGAVITISKITOKLDELKNEELKEKIDTAOKCSEFTNK 120
 Oy 123 SEHAVLGIDNLTDDNAORAILKHKANKGAAELEKLFKAVENTLSKAQDTLKNVKE 182
 DB 121 GEHAGLGIENVTDENAKKAILITDADKGAELEKLFKAVENTLSKAKEMLSNVKE 180
 Oy 183 SPV 186
 DB 181 SPV 184

RESULT 13
 ID Q9R7B1 PRELIMINARY: PRT: 192 AA.
 AC Q9R7B1: 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
 GN OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxId-139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TESTS;
 RA MEDLINE-97478003; PubMed-9336916;
 RA Ras N.M., Postic D., Foretz M., Baranton G.;
 "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
 U.S.A.'?"
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TESTS;
 RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
 Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U91798; AAB81895.1;-
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT SEQUENCE 192 AA: 20297 MW: 6770502A20AAAF64 CRC64;

Query Match 34.7%; Score 623.5; DB 2; Length 192;
 Best Local Similarity 69.5%; Pred. No. 2.3e-22;
 Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

Oy 1 MACNSGKDGNAASNSADSVKGNLTETISKITESNAVLAKEVEETLLASIDELATKA 60
 DB 3 ISCNSSKDKDNTSNAASNSADSVKGNLTETISKITESNAVLAKEVEETLLASIDELATKA 62
 Oy 61 IGKRI-GNNGLEANQKNTSILSGAVAISDLIAEKLNVKNEELKEKIDTAOKCSTFTNK 119

```

Db      63  IGKIHQNNGLDTENNHNGSLAGAVAIISLILNKDGLKNEGLKEDIKIDAKKCSSEFTN 122
QY      120  KLKSEHAVLGLDNLTDNNAORAILKKHANKDGAELKLFKAVENLSKAAOFTLKNAYK 179
Db      123  KLEKHTDLGKEGVTDADAKAELIKTNGTKGAELIKLFESVEVLSKAAKEMLANSVK 182
QY      180  ELTSPV 186
Db      183  ELTSPV 189

RESULT 14
Q9RR53  PRELIMINARY; PRT; 193 AA.
AC      09RR53;
DT      01-MAY-2000 (TREMblrel. 13, Created)
DT      01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT      01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C (FRAGMENT).
GN      OSCP.
OS      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID-139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-OC1.
RX      MEDLINE-99091544; PubMed-9872945;
RA      Wang J.N., Dykhuizen D.E., Olu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT      "Genetic diversity of oSCP in a local population of Borrelia
RL      burgdorferi sensu stricto."
RL      Genetics 151:15-30(1999).
DR      EMBL; AF029860; AAB86543.1;
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
FT      NON_TER 1
FT      NON_TER 193
SQ      SEQUENCE 193 AA; 20502 MW; 5EFD5AF8986D1E CRC64;

```

```

Query Match      34.7%; Score 623.5; DB 2; Length 193;
Best Local Similarity 69.5%; Pred. No. 2.3e-22;
Matches 130; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

```

```

QY      1  MACNNSGDKGNASANSADSVKGPMLTEISKRTESNAVLAKEVEFTLASIDELATKA 60
       7  ISCNNSGDKGNTSANSADSVKGPMLTEISKRTDSNAVLAVKEVEFTLASIDELATKA 66
QY      61  IGKKT-IGNNGLEANOSKNTSLSGAVAIISDLIAEKLNVLAKEVEFTLASIDELATKA 119
Db      67  IGKTHQNNGLDTENNHNGSLAGAVAIISLILNKDGLKNEGLKEDIKIDAKKCSSEFTN 126
QY      120  KLKSEHAVLGLDNLTDNNAORAILKKHANKDGAELKLFKAVENLSKAAOFTLKNAYK 179
Db      127  KLEKHTDLGKEGVTDADAKAELIKTNGTKGAELIKLFESVEVLSKAAKEMLANSVK 186
QY      180  ELTSPV 186
Db      187  ELTSPV 189

RESULT 15
Q31123  PRELIMINARY; PRT; 185 AA.
AC      Q31123;
DT      01-JAN-1998 (TREMblrel. 05, Created)
DT      01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT      01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C (FRAGMENT).
GN      OSCP.
OS      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

```

```

OX      NCBI_TaxID-139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-OC13;
RA      Wang J.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL      Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF029872; AAB86555.1;
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
FT      NON_TER 1
FT      NON_TER 185
SQ      SEQUENCE 185 AA; 19673 MW; 58D6FEE3C7769CAF CRC64;

```

```

Query Match      34.7%; Score 623; DB 2; Length 185;
Best Local Similarity 72.1%; Pred. No. 2.3e-22;
Matches 129; Conservative 24; Mismatches 24; Indels 2; Gaps 2;

```

```

QY      1  MACNNSGDKGNASANSADSVKGPMLTEISKRTESNAVLAKEVEFTLASIDELATKA 60
       7  ISCNNSGDKGNTSANSADSVKGPMLTEISKRTESNAVLAKEVEFTLASIDELATKA 66
QY      61  IGKKT-IGNNGLEANOSKNTSLSGAVAIISDLIAEKLNVLAKEVEFTLASIDELATKA 118
Db      67  IGKTHQNNGLDTENNHNGSLAGAVAIISLILNKDGLKNEGLKEDIKIDAKKCSSEFTN 126
QY      119  KLKSEHAVLGLDNLTDNNAORAILKKHANKDGAELKLFKAVENLSKAAOFTLKNAY 177
Db      127  KLEGEHAGLGIENVTDENNAKAILITDAKDGAELKLFKAVENLSKAAKEMLANSVK 185

```

```

Search completed: March 18, 2002, 10:10:55
Job time: 978 sec

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:33 ; Search time 118.14 Seconds

230.734 Million cell updates/sec

Title: US-09-596-746A-42

Sequence: 1 MACNNSGKDGNASANSADES.....AVENLAKAKAKEMILANSVKEL 368

Scoring table: BLOSUM62

Searched: 522463 seqs, 74073290 residues

total number of hits satisfying chosen parameters: 522463

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_1101:*

1:	/S1D58/gcgdata/geneseq/geneseqp/AA1960.DAT *
2:	/S1D58/gcgdata/geneseq/geneseqp/AA1981.DAT *
3:	/S1D58/gcgdata/geneseq/geneseqp/AA1982.DAT *
4:	/S1D58/gcgdata/geneseq/geneseqp/AA1983.DAT *
5:	/S1D58/gcgdata/geneseq/geneseqp/AA1984.DAT *
6:	/S1D58/gcgdata/geneseq/geneseqp/AA1985.DAT *
7:	/S1D58/gcgdata/geneseq/geneseqp/AA1986.DAT *
8:	/S1D58/gcgdata/geneseq/geneseqp/AA1987.DAT *
9:	/S1D58/gcgdata/geneseq/geneseqp/AA1988.DAT *
10:	/S1D58/gcgdata/geneseq/geneseqp/AA1989.DAT *
11:	/S1D58/gcgdata/geneseq/geneseqp/AA1990.DAT *
12:	/S1D58/gcgdata/geneseq/geneseqp/AA1991.DAT *
13:	/S1D58/gcgdata/geneseq/geneseqp/AA1992.DAT *
14:	/S1D58/gcgdata/geneseq/geneseqp/AA1993.DAT *
15:	/S1D58/gcgdata/geneseq/geneseqp/AA1994.DAT *
16:	/S1D58/gcgdata/geneseq/geneseqp/AA1995.DAT *
17:	/S1D58/gcgdata/geneseq/geneseqp/AA1996.DAT *
18:	/S1D58/gcgdata/geneseq/geneseqp/AA1997.DAT *
19:	/S1D58/gcgdata/geneseq/geneseqp/AA1998.DAT *
20:	/S1D58/gcgdata/geneseq/geneseqp/AA1999.DAT *
21:	/S1D58/gcgdata/geneseq/geneseqp/AA2000.DAT *
22:	/S1D58/gcgdata/geneseq/geneseqp/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	ID	Description
			Match	Length	
	1	1793	99.7	367	22 AAB65719
	2	1790	99.6	391	22 AAB65735
	3	1557.5	86.6	368	22 AAB65718
	4	1554.5	86.5	392	22 AAB65734
	5	1538	85.5	368	22 AAB65717
	6	1530	85.1	391	22 AAB65736
	7	1485.5	82.6	377	22 AAB65713
	8	1484	82.5	373	22 AAB65711
	9	1481	82.4	397	22 AAB65729
	10	1480	82.3	369	22 AAB65716
	11	1479.5	82.3	400	22 AAB65739

[illegible]

	RESULT	ALIGNMENTS
PT	1	
TT	AAB62719	
CT	AAB62719 standard; Protein; 367 AA.	
GT	'AC	
AT	AAB62719;	
DT	03-APR-2001 (first entry)	
DE	Borrelia sp chimeric ospc protein SEQ ID NO: 42.	
XX		
KW	Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.	
XX		
OS	Chimeric - Borrelia sp.	
OS	Chimeric - Borrelia sp.	
XX		
PN	MO200078936-A1.	
PD		
PD	28-DEC-2000.	
PF	19-JUN-2000; 2000MO-US16915.	
XX		
PR	18-JUN-1999; 99US-0140042.	
XX		
PA	(UANY) UNIV NEW YORK STATE RES FOUND.	
XX	(BROO-) BROOK BIOTECHNOLOGIES INC.	
PI	Datwyler RJ, Seino G, Dykhuzen D, Luft BJ, Gomes-Solecki M;	
XX		
XX	WPI; 2001-050113/06.	
DR	N-PSDB; AAF29023.	
PT	Compositions of OspC polypeptides from strains of Borrelia which causes	
TT	Lyme disease are used to immunize animals and detect immune responses	
CT	to Lyme disease -	

XX Claim 43: Page 102-103; 160pp; English.
 PS
 XX
 CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 CC
 SQ Sequence 367 AA:

Query Match 99.7%; Score 1793; DB 22; Length 367;
 Best Local Similarity 100.0%; Pred. No. 1.2e-108;
 Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGMSANSADSVKGPNTLTKITSSNAVLAKEVEETLLASIDELATKAI 61
 1 acnnsqkdgmsansadesvkgpnlteiskkltesnavlavkevelasidelatka1 60
 QY 62 GKTIQNGCLEANQSKNTSLSGAVAIISDLIAEKLVNLEKNELEKIDTAQOCSTEFNKL 121
 61 gkklqngcleangskntslsgayaisdliaklnvlnkeelkeidtaqocstefnkl 120
 QY 122 KSEHAVIGLDLUTDDNQRALIKKANKDKGALEKLFRAVENLSKRAODTLKNAVKEL 181
 121 ksehavlgldluttddnqralikkhankdkgaaleklfravenlskaagdtlknavekl 180
 QY 182 TSPIVHGNSGKDGNTSANSADSVKGPNTLTKITSSNAVLAKEVEETLLASIDEL 241
 181 tspivhgnsqkdgntsansadesvkgpnlteiskkltesnavlavkeletllasidel 240
 QY 242 ATKAIGKKIQOONGLAVEACHNGTLLAGAYTISKLTITQKIDGKNSKLEKIEKAKKCS 301
 241 atkaigkqiqngglaveachngtllagaytiskltitqkldgknsklekienakcs 300
 QY 302 EDFTKLGEHQAQIGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAKKEML 361
 301 edftkklgehaqigienvtdenakkailitdaakdkgaaleklfkavenlakaakeml 360
 Db 361 ANSVKEL 368
 361 ansvkel 367

RESULT 2
 AAB62735

AAB62735 standard; Protein; 391 AA.

AC AAB62735;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 74.

KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYV) UNIV NEW YORK STATE RES FOUND.
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.
 PI Dattwyler RJ, Selmost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
 XX

DR WPI: 2001-050113/06.
 DR N-PDB; AAF29039.
 XX
 XX
 PT Compositions of ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX
 PS Claim 43; Page 142; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 CC
 SQ Sequence 391 AA:

Query Match 99.6%; Score 1790; DB 22; Length 391;
 Best Local Similarity 99.7%; Pred. No. 2e-108;
 Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDGMSANSADSVKGPNTLTKITSSNAVLAKEVEETLLASIDELATKAI 61
 25 acnnsqkdgmsansadesvkgpnlteiskkltesnavlavkevelasidelatka1 64
 QY 62 GKTIQNGCLEANQSKNTSLSGAVAIISDLIAEKLVNLEKNELEKIDTAQOCSTEFNKL 121
 85 gkklqngcleangskntslsgayaisdliaklnvlnkeelkeidtaqocstefnkl 114
 QY 122 KSEHAVIGLDLUTDDNQRALIKKANKDKGALEKLFRAVENLSKRAODTLKNAVKEL 181
 145 ksehavlgldluttddnqralikkhankdkgaaleklfravenlskaagdtlknavekl 204
 QY 182 TSPIVHGNSGKDGNTSANSADSVKGPNTLTKITSSNAVLAKEVEETLLASIDEL 241
 205 tspivhgnsqkdgntsansadesvkgpnlteiskkltesnavlavkeletllasidel 264
 QY 242 ATKAIGKKIQOONGLAVEACHNGTLLAGAYTISKLTITQKIDGKNSKLEKIEKAKKCS 301
 265 atkaigkqiqngglaveachngtllagaytiskltitqkldgknsklekienakcs 324
 QY 302 EDFTKLGEHQAQIGIENVTDENAKKAILITDAKDKGALEKLFKAVENLAKAKKEML 361
 325 edftkklgehaqigienvtdenakkailitdaakdkgaaleklfkavenlakaakeml 384
 Db 362 ANSVKEL 368
 362 ansvkel 391

RESULT 3

AAB62718 standard; Protein; 368 AA.

ID AAB62718;

AC AAB62718;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 40.

KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.

OS Chimeric - Borrelia sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

XX (UANY) UNTV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI: 2001-050113/06.
DR N-PSDB: AAP29022.
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43; Page 99-100; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

Sequence 368 AA;

Query Match 86.6%; Score 1557.5; DB 22; Length 368;
Best Local Similarity 87.5%; Pred. No. 2.1e-93;
Matches 321; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 2 ACNNSGKDNASANSADSVKGNLTETISKRTESNAVVLAVKEVETLLASIDELATKAI 61
DB 1 acnngskdgnasansadesvkgpnltetiskrtesnavvlavkevetllasidelatka 60
OY 62 GKRIKGNGLANOSKNTSLSGAVAISDLIAEKLNVLKNELKEKIDTAKOCSTETFNKL 121
DB 61 gkrikgngleangskntsllsgavaisdliakelnvlnkeelkekldtakqscstefnkl 120
OY 122 KSEHNAVIGLNDITDDNAQRALIKKHANKDGALEKLFRAVENLSKAADDTLKNAYKEL 181
DB 121 ksehnavlgldntddnaqrallkhnankdgaaleklfravenlskaadqtlknaykel 180
OY 182 TSPIVHNGNSGKDNSTANSADSVKGNLTETISKRTESNAVVLAVKEVETLLASIDEL 241
DB 181 tspivhngnsgkdntsansadesvkgpnltetiskrtesnavvlavkevetllasidel 240
OY 242 ATRAIKGIKIQONGLAVERHNGTLLAGAYTISKLTOKLDGKNSKLEKIEENAKKCS 301
DB 241 a-kaigkikindvsldeahngslisgylisnlltkksaikdsgeklkaelekakcs 299
OY 302 EDTFTKLEGEHAOLGIENTVDENAKKAILITDAKDKGALEKLFRAVENLAKAAKEML 361
DB 300 eedftaklgehtldgkyevtdnnaakailtkndcktgadaleklfesvnlkskaakeml 359
OY 362 ANSVKEL 368
DB 360 tmsvkel 366

RESULT 4
AAB62734
ID AAB62734 standard; Protein; 392 AA.
XX
XX AAB62734;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 72.
XX
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
XX WO200078966-A1.

XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000WO-US16915.
XX
XX 18-JUN-1999; 99US-0140042.
XX
XX (UANY) UNTV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX Datwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX WPI: 2001-050113/06.
DR N-PSDB: AAP29038.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43; Page 139-140; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

Sequence 392 AA;

Query Match 86.5%; Score 1554.5; DB 22; Length 392;
Best Local Similarity 87.2%; Pred. No. 3.5e-93;
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

OY 2 ACNNSGKDNASANSADSVKGNLTETISKRTESNAVVLAVKEVETLLASIDELATKAI 61
DB 25 acnngskdgnasansadesvkgpnltetiskrtesnavvlavkevetllasidelatka 84
OY 62 GKRIKGNGLANOSKNTSLSGAVAISDLIAEKLNVLKNELKEKIDTAKOCSTETFNKL 121
DB 85 gkrikgngleangskntsllsgavaisdliakelnvlnkeelkekldtakqscstefnkl 144
OY 122 KSEHNAVIGLNDITDDNAQRALIKKHANKDGALEKLFRAVENLSKAADDTLKNAYKEL 181
DB 145 ksehnavlgldntddnaqrallkhnankdgaaleklfravenlskaadqtlknaykel 204
OY 182 TSPIVHNGNSGKDNSTANSADSVKGNLTETISKRTESNAVVLAVKEVETLLASIDEL 241
DB 205 tspivhngnsgkdntsansadesvkgpnltetiskrtesnavvlavkevetllasidel 264
OY 242 ATRAIKGIKIQONGLAVERHNGTLLAGAYTISKLTOKLDGKNSKLEKIEENAKKCS 301
DB 265 a-kaigkikindvsldeahngslisgylisnlltkksaikdsgeklkaelekakcs 323
OY 302 EDTFTKLEGEHAOLGIENTVDENAKKAILITDAKDKGALEKLFRAVENLAKAAKEML 361
DB 324 eedftaklgehtldgkyevtdnnaakailtkndcktgadaleklfesvnlkskaakeml 383
OY 362 ANSVKEL 368
DB 384 tmsvkel 390

RESULT 5
AAB62717
ID AAB62717 standard; Protein; 368 AA.
XX
XX AAB62717;
XX
XX 03-APR-2001 (first entry)
XX
XX Borrelia sp chimeric ospC protein SEQ ID NO: 38.

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UANY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
N-PSDB; AAF29021.
XX
XX WPI; 2001-050113/06.
XX
XX N-PSDB; AAF29021.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 97; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
SQ Sequence 368 AA:

Query Match 85.5%; Score 1538; DB 22; Length 368;
Best Local Similarity 86.7%; Pred. No. 3.8e-92;
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

QY 1 MACNSGKDGMSANSADSVKGNPLTEISKRTTESNAVAVLAVKEVETLLASIDELATKA 60
DB 1 macnsgkdgnasansadesvkgpnlteiskrttesnavlavkevettllasidelatka 60
QY 61 ICKKIGNNGLEANSKNTSLSGAYASDLAEKLVNLEKLEKIDTAKOCSTEFNTK 120
DB 61 ickkignngleangskntslsgayaisdlaeeklvnlkneelkeidtakgcsstefntk 120
QY 121 LKSEHAVLGLDNLTDNAORAILTKHANKDKGAELKLFKAVENLSRAAODTLKNAVKEL 180
DB 121 lksehavlgldnltdnagrailtkhankdkgaaleklfkavenlskaagdtlknavek 180
QY 181 LTPSPVHGNSGKDGNTSANSADSVKGNPLTEISKRTTESNAVAVLAVKEVETLLASIDE 240
DB 181 ltpspvhgnsrkdgnasnsadesvkgpnlteiskrttesnavlavkevettllaside 240
QY 241 LATKAIIGKRIQONGLAVEAGHNGTLGAYTISKLTIQKLDGLKNSKLEKTEENAKKC 300
DB 241 latkaiigkriqonglavaghnngtllagaytiskltiokldglknskleteenakkc 300
QY 301 SEDFTKLEGEHAQIGIENVTDENAKKAILITDAKDKGAELKLFKAVENLSKAADTL 360
DB 301 sedftklegehaolgienvtdenakkailittdakdkgaaleklfkavenlskaagdtl 360
QY 361 LANSVKEL 368
DB 361 lnavkel 366
RESULT 6
ID AAB62736 standard; Protein; 391 AA.
XX

AC AAB62736;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 76.
XX
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UANY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
N-PSDB; AAF29040.
XX
XX WPI; 2001-050113/06.
XX
XX N-PSDB; AAF29040.
XX
XX Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 144-145; 160pp; English.
XX
XX The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
XX to Lyme disease.
XX
SQ Sequence 391 AA:

Query Match 85.1%; Score 1530; DB 22; Length 391;
Best Local Similarity 86.4%; Pred. No. 1.4e-91;
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

QY 2 ACNNSGKGNMSANSADSVKGNPLTEISKRTTESNAVAVLAVKEVETLLASIDELATKA 61
DB 2 acnnsdgkgnasansadesvkgpnlteiskrttesnavlavkevettllasidelatka 61
QY 62 GKKTGNNGLEANSKNTSLSGAYASDLAEKLVNLEKLEKIDTAKOCSTEFNTK 121
DB 62 gkktgnngleangskntslsgayaisdlaeeklvnlkneelkeidtakgcsstefntk 121
QY 144 KSEHAVLGLDNLTDNAORAILTKHANKDKGAELKLFKAVENLSRAAODTLKNAVKEL 181
DB 144 ksehavlgldnltdnagrailtkhankdkgaaleklfkavenlskaagdtlknavek 181
QY 182 TSPVHGNSGKDGNTSANSADSVKGNPLTEISKRTTESNAVAVLAVKEVETLLASIDE 241
DB 182 tspvhgnsrkdgnasnsadesvkgpnlteiskrttesnavlavkevettllaside 241
QY 242 ATKAIIGKRIQONGLAVEAGHNGTLGAYTISKLTIQKLDGLKNSKLEKTEENAKKC 301
DB 242 atkaiigkriqonglavaghnngtllagaytiskltiokldglknskleteenakkc 301
QY 302 EDFTKLEGEHAQIGIENVTDENAKKAILITDAKDKGAELKLFKAVENLSKAADTL 361
DB 302 edftklegehaolgienvtdenakkailittdakdkgaaleklfkavenlskaagdtl 361
QY 362 ANSVKEL 368
DB 362 knavek 389

QY 236 ASIDELATKAIGKRIQONGGLAVEAGHNGTLLAGAVYITISKLITOKLDGLKNSKLEKIE 295
Db 241 asidelatkaigkriqngglaveaghgntllagayitisklittqkldglknskilekie 300
QY 296 NAKKSEDEFTFKLGEHAQOLCIENVTDENAKKAILITDAKDGALEKLFKAVENIAK 355
Db 301 nakkosedftkklgehaqgienvtdenakkailitdaakdgaaelekifkaveniak 360
QY 356 AAKEMIANSVKEL 368
Db 361 aakemiansvkel 373

RESULT 9
AAB62729
ID AAB62729 standard; Protein: 397 AA.

QY AAB62729;
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SFQ ID NO: 62.
XX
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29033.
XX
PT Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 125-126; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 397 AA;

Query Match 82.4%; Score 1481; DB 22; Length 397;
Best Local Similarity 83.1%; Pred. No. 2e-88;
Matches 310; Conservative 23; Mismatches 34; Indels 6; Gaps 2;

QY 2 ACNNSGKDGANSANSADSVKGPMLTEISKITSNVAVLAKVEVTLASIDELATKAI 61
Db 25 scnmsgdgntsansadesvkgpmlteiskitdsnavllavkeveallsideiaakai 84
QY 62 GKRI-GNNGLEANOSKMTSLISGAYAISDLAEKLVNKLNELEKIDTAOCSTEFNK 120
Db 85 gkrlhngngldteynhngslilagayaisclikqkldglkneglekidaakccsetlntk 144
QY 121 LKSHAVALGIDNLTDNAQRALTKKHNKDKGAELKLFKAVENLSKAQDTLKNVKE 180

Db 145 lkehtldygkegyvdadakeaalktngltkgyaeelglfsevevlskaakemiansvke 204
QY 181 LTSPLVHG-----NNSGKDGNTSANSADSVKGPMLTEISKITSNVAVLAKVEITEL 235
Db 205 ltsplvhaespmvnhsgdgntsansadesvkgpmlteiskitdsnavllavkelell 264
QY 236 ASIDELATKAIGKRIQONGGLAVEAGHNGTLLAGAVYITISKLITOKLDGLKNSKLEKIE 295
Db 265 asidelatkaigkriqngglaveaghgntllagayitisklittqkldglknskilekie 324
QY 296 NAKKSEDEFTFKLGEHAQOLCIENVTDENAKKAILITDAKDGALEKLFKAVENIAK 355
Db 325 nakkosedftkklgehaqgienvtdenakkailitdaakdgaaelekifkaveniak 384
QY 356 AAKEMIANSVKEL 368
Db 385 aakemiansvkel 397

RESULT 10
AAB62716
ID AAB62716 standard; Protein: 369 AA.

QY AAB62716;
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SFQ ID NO: 36.
XX
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29020.
XX
PT Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 94-95; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 369 AA;

Query Match 82.3%; Score 1480; DB 22; Length 369;
Best Local Similarity 83.8%; Pred. No. 2.2e-88;
Matches 310; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDGANSANSADSVKGPMLTEISKITSNVAVLAKVEVTLASIDELATKAI 61
Db 1 acnmsgdgntsansadesvkgpmlteiskitdsnavllavkeveallsideiaakai 59

CC to Lyme disease.

XX Sequence 393 AA;

Query Match 82.1%; Score 1477; DB 22; Length 393;
 Best Local Similarity 83.5%; Pred. No. 3.7e-88;
 Matches 309; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

QY 2 ACNNSGKDGNSANSADSVKGNLTETISKRITESNAVLAWEVETLLASIDELATKA 61
 Db 25 scnsqskdgnstansadesvkgpnlteiskritesnavlavkeveallssidela-kai 83
 QY 62 GKRTGNGG-LEANQSKNTSLSGAVATSDLAETKLVLANEE-LKEKIDTAKOCSTEF 119
 Db 84 gkktkhdgnsidneanrnesltagaystltctqklskngseglkekiakkcseft 143
 QY 120 KIKSEHAVIGLDLTDNNQRAILKRN-ANKDKGAELKEKFAVENLSKAADTLKANV 178
 Db 144 ktkdmhaqigqvidenakkalkanaagkdkgyeeleklsqslsleslkaakemlansv 203
 QY 179 KELTSPVHGNNSGKDGNTSANSADSVKGNLTETISKRITESNAVLAWEVETLLASI 238
 Db 204 keltspvvhgnmsgkdgntansadesvkgpnlteiskritesnavlavkeveallssidela-kai 263
 QY 239 DELATATIKKTIQONGCLAVEAGHNGTLLAGAVTTSKLTORLIDGLKNSKLEKIE 298
 Db 264 delatkaiktkiqngglaveaghnqgllagaystltctqklskngseglkekiakkcseft 323
 QY 299 KCSQEDTKLEGEHQAOLGIENTVDENAKAAILITDAKDKGAELKEKFAVENLSKA 358
 Db 324 kcsedtkklegehaglgienvdenakkailitdaakdgaaelekfkavenlakaak 383
 QY 359 EMLANSVKEL 368
 Db 384 emlansvkel 393

RESULT 13

AAB62712 ID AAB62712 standard; Protein: 378 AA.

XX AAB62712;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 28.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

XX N-PSDB; AAF29016.

XX Compositions of OspC polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 PT to Lyme disease -
 XX

PS Claim 43; Page 83-84; 160pp; English.

XX The present invention provides compositions comprising ospC proteins and
 CC chimeric ospC proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.
 XX

SO Sequence 378 AA;

Query Match 69.5%; Score 1250; DB 22; Length 378;
 Best Local Similarity 70.3%; Pred. No. 1.7e-73;
 Matches 265; Conservative 43; Mismatches 59; Indels 10; Gaps 3;

QY 1 MACNNSGKDGNSANSADSVKGNLTETISKRITESNAVLAWEVETLLASIDELATKA 60
 Db 1 macnnsqkdgntansadesvkgpnlteiskritesnavlavkeveallssidelaika 60
 QY 61 IGKRT-GNNGLEANQSKNTSLSGAVATSDLAETKLVLANEE-LKEKIDTAKOCSTEF 119
 Db 61 igkrtgngldtneahngslagayaistltqkldgkngelkekiakkcseftn 120
 QY 120 KIKSEHAVIGLDLTDNAQRAILKRN-ANKDKGAELKEKFAVENLSKAADTLKANV 179
 Db 121 ktkdmhaqigqvidenakkalkanaagkdkgyeeleklsqslsleslkaakemlansv 180
 QY 180 ELTSPVHGNNSGKDGNTSANSADSVKGNLTETISKRITESNAVLAWEVETLLASI 231
 Db 181 eltspvvhgnmsgkdgntansadesvkgpnlteiskritesnavlavkeveallssidela-kai 240
 QY 232 ETLASIDELATKAIGKRTIQONGCLAVEAGHNGTLLAGAVTTSKLTORLIDGLKNSK 291
 Db 241 etlltsidela-kaiqkktkhdgnsidneanrnesltagaystltctqklskngseglkekiakkcseft 299
 QY 292 EKIEKAKCSEDTTKLEGEHQAOLGIENTVDENAKAAILITDAKDKGAELKEKFAVE 351
 Db 300 aeiekaakcseefaklkgehtldgkegyvddnakkalktkmndktgadelkelfesvk 359
 QY 352 NLAKAEMLANSVKEL 368
 Db 360 nlskaakemlansvkel 376

RESULT 14

AAB62710 ID AAB62710 standard; Protein: 374 AA.

XX AAB62710;

DT 03-APR-2001 (first entry)

XX Borrelia sp chimeric ospC protein SEQ ID NO: 24.

XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.

XX WO200078966-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000WO-US16915.

XX 18-JUN-1999; 99US-0140042.

XX (UYNY) UNIV NEW YORK STATE RES FOUND.

XX (BROO-) BROOK BIOTECHNOLOGIES INC.

XX Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX WPI; 2001-050113/06.

DR N-PSDB: AAF29014.

XX Compositions of Ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 to Lyme disease -

PS Claim 43; Page 78-79; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

SQ Sequence 374 AA;

Query Match 69.4%; Score 1248.5; DB 22; Length 374;
 Best Local Similarity 71.0%; Pred. No. 2.1e-73;
 Matches 265; Conservative 42; Mismatches 59; Indels 7; Gaps 3;

2 ACNNSGKDGNSANSADSVKGPNTLEISKTTESNAVVLAVKEVETLLASIDELATKAI 61
 1 acnsgkdgtansadesvkgpntleiskttidnavllavkeveallssidelaakal 60
 62 GKKT-GNNGLEANSKNTSLSGAVASIDIAEKLVKNEELKEKIDPAKOCSTFETNK 120
 61 gkklhngldteyhnngsllegayalstlkqldgikneglkeidaakkcsetfkn 120
 121 LKSEHAVLGLDNLTDNAORAILKHKANKDKGALEKLFKAVENTSKAAOTLKNAVKE 180
 121 lkehtldlgyegvtdadeaalkngtktgaeelgfesvevlskaakemlansvke 180
 181 LSPPIYHG-----NNSGKDGNTSANSADSVKGPNTLEISKTTESNAVVLAVKEIETLL 235
 181 lspvvaesparwnngskdgtansadesvkgpntleiskttidnavllavkeveell 240
 236 ASTDELATRAIKKIQONGSLAVEAGHNGTLLAGATYISKLTQKIDGLKNSKLEKIE 295
 241 tsidela-kalgkklkndvslneadngslisgaylslntlkksaalkdsgelekaele 299
 296 NAKKCEDFTKKLEGEHAOLGIENVTDENAKKAILTTDAKDKGALEKLFKAVENTLAK 355
 300 kakkceeftaklgehtdlygkegyvddnakkalkltmndktgadelleklfesvknlsk 359
 356 AAKEMLANSVKEL 368
 360 aakemltnsvkel 372

ULF 15

ID AAB62728 standard; Protein: 398 AA.

AC AAB62728;

DT 03-APR-2001 (first entry)

DE Borrelia sp chimeric ospc protein SEQ ID NO: 60.

KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

OS Chimeric - Borrelia sp.
 OS Chimeric - Borrelia sp.

PN W0200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000MO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UNYK) UNIV NEW YORK STATE RES FOUND.

PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyler RJ, Seino G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB: AAF29032.

PT Compositions of Ospc polypeptides from strains of Borrelia which cause
 PT Lyme disease are used to immunize animals and detect immune responses
 to Lyme disease -

PS Claim 43; Page 123; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
 CC chimeric ospc proteins from members of the Borrelia genus. These may be
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
 CC vaccines against Borrelia infection, which is spread by ticks and leads
 CC to Lyme disease.

SQ Sequence 398 AA;

Query Match 69.3%; Score 1245.5; DB 22; Length 398;
 Best Local Similarity 70.8%; Pred. No. 3.6e-73;
 Matches 264; Conservative 43; Mismatches 59; Indels 7; Gaps 3;

2 ACNNSGKDGNSANSADSVKGPNTLEISKTTESNAVVLAVKEVETLLASIDELATKAI 61
 25 acnsgkdgtansadesvkgpntleiskttidnavllavkeveallssidelaakal 84
 62 GKKT-GNNGLEANSKNTSLSGAVASIDIAEKLVKNEELKEKIDPAKOCSTFETNK 120
 62 gkklhngldteyhnngsllegayalstlkqldgikneglkeidaakkcsetfkn 144
 85 gkklhngldteyhnngsllegayalstlkqldgikneglkeidaakkcsetfkn 144
 121 LKSEHAVLGLDNLTDNAORAILKHKANKDKGALEKLFKAVENTSKAAOTLKNAVKE 180
 145 lkehtldlgyegvtdadeaalkngtktgaeelgfesvevlskaakemlansvke 204
 181 LSPPIYHG-----NNSGKDGNTSANSADSVKGPNTLEISKTTESNAVVLAVKEIETLL 235
 181 lspvvaesparwnngskdgtansadesvkgpntleiskttidnavllavkeveell 264
 205 ltpvvaesparwnngskdgtansadesvkgpntleiskttidnavllavkeveell 264
 236 ASTDELATRAIKKIQONGSLAVEAGHNGTLLAGATYISKLTQKIDGLKNSKLEKIE 295
 265 tsidela-kalgkklkndvslneadngslisgaylslntlkksaalkdsgelekaele 323
 296 NAKKCEDFTKKLEGEHAOLGIENVTDENAKKAILTTDAKDKGALEKLFKAVENTLAK 355
 324 kakkceeftaklgehtdlygkegyvddnakkalkltmndktgadelleklfesvknlsk 383
 356 AAKEMLANSVKEL 368
 384 aakemltnsvkel 396

Search completed: March 18, 2002, 09:54:34
 Job time: 332 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:45 ; Search time 621.2 Seconds

(without alignments)
164,485 Million cell updates/sec

Title: US-09-596-746A-42

1798

Sequence: 1 MACNNSGKDGASNSADES.....AVENLAKAKMLANSVKEL 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Pending Patents, AA, Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1798	100.0	368	19	US-09-596-746A-42
2	1793	99.7	367	19	US-09-596-746A-42
3	1790	99.6	391	19	US-09-596-746A-74
4	1790	99.6	392	19	US-09-596-746A-74
5	1562.5	86.9	369	19	US-09-596-746A-40
6	1557.5	86.6	368	19	US-09-596-746A-40
7	1554.5	86.5	392	19	US-09-596-746-72
8	1554.5	86.5	393	19	US-09-596-746A-72
9	1538	85.5	368	19	US-09-596-746-38

10	1538	85.5	368	19	US-09-596-746A-38	Sequence 38, Appl
11	1530	85.1	391	19	US-09-596-746-76	Sequence 76, Appl
12	1530	85.1	392	19	US-09-596-746A-76	Sequence 76, Appl
13	1489	82.8	374	19	US-09-596-746A-26	Sequence 26, Appl
14	1485.5	82.6	377	19	US-09-596-746-30	Sequence 30, Appl
15	1485.5	82.6	377	19	US-09-596-746A-30	Sequence 30, Appl
16	1485	82.6	370	19	US-09-596-746A-36	Sequence 36, Appl
17	1484	82.5	373	19	US-09-596-746-26	Sequence 26, Appl
18	1481	82.4	397	19	US-09-596-746-62	Sequence 62, Appl
19	1481	82.4	398	19	US-09-596-746A-62	Sequence 62, Appl
20	1480	82.3	369	19	US-09-596-746-36	Sequence 36, Appl
21	1479.5	82.3	400	19	US-09-596-746-82	Sequence 82, Appl
22	1479.5	82.3	401	19	US-09-596-746A-82	Sequence 82, Appl
23	1477	82.1	393	19	US-09-596-746-66	Sequence 66, Appl
24	1477	82.1	394	19	US-09-596-746A-66	Sequence 66, Appl
25	1253.5	69.7	375	19	US-09-596-746A-24	Sequence 24, Appl
26	1250	69.5	378	19	US-09-596-746-28	Sequence 28, Appl
27	1250	69.5	378	19	US-09-596-746A-28	Sequence 28, Appl
28	1249.5	69.5	371	19	US-09-596-746A-34	Sequence 34, Appl
29	1248.5	69.4	374	19	US-09-596-746-24	Sequence 24, Appl
30	1245.5	69.3	398	19	US-09-596-746-60	Sequence 60, Appl
31	1245.5	69.3	399	19	US-09-596-746A-60	Sequence 60, Appl
32	1244.5	69.2	370	19	US-09-596-746-34	Sequence 34, Appl
33	1244	69.2	401	19	US-09-596-746A-80	Sequence 80, Appl
34	1244	69.2	402	19	US-09-596-746A-80	Sequence 80, Appl
35	1241.5	69.0	394	19	US-09-596-746-64	Sequence 64, Appl
36	1241.5	69.0	395	19	US-09-596-746A-64	Sequence 64, Appl
37	1225	68.1	370	19	US-09-596-746A-32	Sequence 32, Appl
38	1220	67.9	369	19	US-09-596-746-32	Sequence 32, Appl
39	1217	67.7	393	19	US-09-596-746-68	Sequence 68, Appl
40	1217	67.7	394	19	US-09-596-746A-68	Sequence 68, Appl
41	1167.5	64.9	360	19	US-09-596-746-52	Sequence 52, Appl
42	1167.5	64.9	360	19	US-09-596-746A-52	Sequence 52, Appl
43	1164.5	64.8	385	19	US-09-596-746A-56	Sequence 56, Appl
44	1159.5	64.5	384	19	US-09-596-746-56	Sequence 56, Appl
45	1156.5	64.3	408	19	US-09-596-746-78	Sequence 78, Appl

ALIGNMENTS

RESULT 1
US-09-596-746A-42
; Sequence 42, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seino, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-42

Query Match 100.0%; Score 1798; DB 19; Length 368;
Best Local Similarity 100.0%; Pred. No. 5, 6e-124;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 MACNNSGKDGASNSADESVKGNLFEISKITTESNAVYLAWEVETLLASIDELATKA 60

Db 1 MACNNSGKDNASANSADSVKGPULTEISKITTESNAVLAWEETLLASIDELATKA 60
QY 61 IGKKGNGLEANSKNTSLSGAYAISDLIAEKLVNLEKEKIDTAQOCSTFTNK 120
Db 61 IGKKGNGLEANSKNTSLSGAYAISDLIAEKLVNLEKEKIDTAQOCSTFTNK 120
QY 121 KSEHAVLGIDNLTDDNAQRAILKHKANKDKGAELKLFKAVENLSKAADTLKNAVKEL 180
Db 121 KSEHAVLGIDNLTDDNAQRAILKHKANKDKGAELKLFKAVENLSKAADTLKNAVKEL 180
QY 181 LTPSVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDE 240
Db 181 LTPSVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDE 240
QY 241 LATKAIGKTIQONGGLAVEAGHNGTLLAGAYTISKLTOTKLDGKNSKLEKEKTEENAKKC 300
Db 241 LATKAIGKTIQONGGLAVEAGHNGTLLAGAYTISKLTOTKLDGKNSKLEKEKTEENAKKC 300
QY 301 SEFTFKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEM 360
Db 301 SEFTFKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEM 360
QY 361 LANSVKEL 368
Db 361 LANSVKEL 368

RESULT 2

US-09-596-746-42
; Sequence 42, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ospc Chimera
US-09-596-746-42

Query Match 99.7%; Score 1793; DB 19; Length 367;
Best Local Similarity 100.0%; Pred. No. 1,3e-123;
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDNASANSADSVKGPULTEISKITTESNAVLAWEETLLASIDELATKA 61
Db 1 ACNNSGKDNASANSADSVKGPULTEISKITTESNAVLAWEETLLASIDELATKA 60
QY 62 GKKGNGLEANSKNTSLSGAYAISDLIAEKLVNLEKEKIDTAQOCSTFTNK 121
Db 62 GKKGNGLEANSKNTSLSGAYAISDLIAEKLVNLEKEKIDTAQOCSTFTNK 120
QY 122 KSEHAVLGIDNLTDDNAQRAILKHKANKDKGAELKLFKAVENLSKAADTLKNAVKEL 181
Db 122 KSEHAVLGIDNLTDDNAQRAILKHKANKDKGAELKLFKAVENLSKAADTLKNAVKEL 180
QY 182 TSPVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDEL 241
Db 182 TSPVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDEL 241

Db 181 TSPVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDEL 240
QY 242 ATRKAIGKTIQONGGLAVEAGHNGTLLAGAYTISKLTOTKLDGKNSKLEKEKTEENAKKCS 301
Db 242 ATRKAIGKTIQONGGLAVEAGHNGTLLAGAYTISKLTOTKLDGKNSKLEKEKTEENAKKCS 300
QY 302 SEFTFKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEM 361
Db 302 SEFTFKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEM 360
QY 362 ANSVKEL 368
Db 362 ANSVKEL 367

RESULT 3

US-09-596-746-74
; Sequence 74, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seimost, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 391
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-74

Query Match 99.6%; Score 1790; DB 19; Length 391;
Best Local Similarity 99.7%; Pred. No. 2,4e-123;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACNNSGKDNASANSADSVKGPULTEISKITTESNAVLAWEETLLASIDELATKA 61
Db 25 SCNNSGKDNASANSADSVKGPULTEISKITTESNAVLAWEETLLASIDELATKA 84
QY 62 GKKGNGLEANSKNTSLSGAYAISDLIAEKLVNLEKEKIDTAQOCSTFTNK 121
Db 85 GKKGNGLEANSKNTSLSGAYAISDLIAEKLVNLEKEKIDTAQOCSTFTNK 144
QY 122 KSEHAVLGIDNLTDDNAQRAILKHKANKDKGAELKLFKAVENLSKAADTLKNAVKEL 181
Db 145 KSEHAVLGIDNLTDDNAQRAILKHKANKDKGAELKLFKAVENLSKAADTLKNAVKEL 204
QY 182 TSPVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDEL 241
Db 205 TSPVHGNNSGKDNNTSANSADSVKGPULTEISKITTESNAVLAWEETLLASIDEL 264
QY 242 ATRKAIGKTIQONGGLAVEAGHNGTLLAGAYTISKLTOTKLDGKNSKLEKEKTEENAKKCS 301
Db 265 ATRKAIGKTIQONGGLAVEAGHNGTLLAGAYTISKLTOTKLDGKNSKLEKEKTEENAKKCS 324
QY 302 SEFTFKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEM 361
Db 325 SEFTFKLEGEHQAOLGIENTVDENAKKAILITDAKDKGAELKLFKAVENLAKAKEM 384
QY 362 ANSVKEL 368
Db 385 ANSVKEL 391

```

RESULT 4
US-09-596-746A-74
; Sequence 74, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoost, Gerald
; APPLICANT: Dykhuisen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIORITY FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 392
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746A-74

Query Match      99.6%; Score 1790; DB 19; Length 392;
Best Local Similarity 99.7%; Pred. No. 2.4e-123;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACNNSGKDNANASADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 61
DB 26 SCNNSGKDNANASADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKAI 85
OY 62 GKKGNGLEANSKNTSLSGAYVATSLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 121
DB 86 GKKGNGLEANSKNTSLSGAYVATSLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 145
OY 122 KSEHAVALGIDNLTDDNAORAILKKHANKDGALELEKLFRAVENLSKAADDTLKNVKE 181
DB 146 KSEHAVALGIDNLTDDNAORAILKKHANKDGALELEKLFRAVENLSKAADDTLKNVKE 205
OY 182 TSPVHGNNSGKDNANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 241
DB 206 TSPVHGNNSGKDNANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 265
OY 242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAVTISKLTQKLDGLKSEKLEKTEMAKKS 301
DB 266 ATKAIGKKIQONGGLAVEAGHNGTLLAGAVTISKLTQKLDGLKSEKLEKTEMAKKS 325
OY 302 EDTTKLEGEHNOIGTENTDENAKKAILITDAKDKGALEKLFRAVENLSKAADDTLKNVKE 361
DB 326 EDTTKLEGEHNOIGTENTDENAKKAILITDAKDKGALEKLFRAVENLSKAADDTLKNVKE 385
OY 362 ANSVKEL 368
DB 386 ANSVKEL 392

```

```

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-40

Query Match      86.9%; Score 1562.5; DB 19; Length 369;
Best Local Similarity 87.5%; Pred. No. 1.3e-106;
Matches 322; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

OY 1 MACNNSGKDNANASADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKA 60
DB 1 MACNNSGKDNANASADESVKGNLTETSKITTESNAVVLAVKEVETLLASIDELATKA 60
OY 61 IGKKGNGLEANSKNTSLSGAYVATSLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 120
DB 61 IGKKGNGLEANSKNTSLSGAYVATSLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 120
OY 121 KSEHAVALGIDNLTDDNAORAILKKHANKDGALELEKLFRAVENLSKAADDTLKNVKE 180
DB 121 KSEHAVALGIDNLTDDNAORAILKKHANKDGALELEKLFRAVENLSKAADDTLKNVKE 180
OY 181 TSPVHGNNSGKDNANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 240
DB 181 TSPVHGNNSGKDNANSADSVKGNLTETSKITTESNAVVLAVKEVETLLASIDEL 240
OY 241 LATKAIGKKIQONGGLAVEAGHNGTLLAGAVTISKLTQKLDGLKSEKLEKTEMAKKS 300
DB 241 LA-KAIGKKIKNDVSLDNEADHNGSLISGAYVATSLIAEKLNVKNEELKEKIDTAKOCSTETFTNK 299
OY 301 SEDFTKLEGEHNOIGTENTDENAKKAILITDAKDKGALEKLFRAVENLSKAADDTLKNVKE 360
DB 300 SEDFTKLEGEHNOIGTENTDENAKKAILITDAKDKGALEKLFRAVENLSKAADDTLKNVKE 359
OY 361 ANSVKEL 368
DB 360 LTNSVKEL 367

RESULT 6
US-09-596-746-40
; Sequence 40, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoost, Gerald
; APPLICANT: Dykhuisen, Daniael
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIORITY FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-40

Query Match      86.6%; Score 1557.5; DB 19; Length 368;

```

Best Local Similarity 87.5%; Pred. No. 3e-106;
Matches 321; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDELATKAI 61
DB 1 ACNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDELATKAI 60
OY 62 GKRIKGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNELEKIDITAKOCSTEFNTKL 121
DB 61 GKRIKGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNELEKIDITAKOCSTEFNTKL 120
OY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKLEKRAVENLSKAAODTLKNAVKE 181
DB 121 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKLEKRAVENLSKAAODTLKNAVKE 180
OY 182 TSPVHGNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDEL 241
DB 181 TSPVHGNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDEL 240
OY 242 ATRAIKRIKIQONGGLAVEGHNGTLLAGAYTTSKLTITKLDKLNSEKLEKIEENAKKS 301
DB 241 A-KAIGKRIKNDVSLDNEADHNGSLISGAYILSNLTITKISAIKDSGELKAEIEKAKKS 299
OY 302 EEDTKLGEHQAQIGENTYDENAKKAILITDAKDKGAELKLEKRAVENLSKAAKEM 361
DB 300 EEDTKLGEHQAQIGENTYDENAKKAILITNDKTKGADELEKLEFESVKLSKAAKEM 359
OY 362 ANSVKEL 368
DB 360 TNSVKEL 366

RESULT 7
US-09-596-746-72
; Sequence 72, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596.746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140.042
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 392
; TYPE: PRN
; ORGANISM: ospc Chimera
US-09-596-746-72
```

Query Match 86.5%; Score 1554.5; DB 19; Length 392;
Best Local Similarity 87.2%; Pred. No. 5.4e-106;
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDELATKAI 61
DB 25 SCNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDELATKAI 84
OY 62 GKRIKGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNELEKIDITAKOCSTEFNTKL 121
DB 85 GKRIKGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNELEKIDITAKOCSTEFNTKL 144
OY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKLEKRAVENLSKAAODTLKNAVKE 181
DB 145 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKLEKRAVENLSKAAODTLKNAVKE 204
```

```
OY 182 TSPVHGNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDEL 241
DB 205 TSPVHGNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDEL 264
OY 242 ATRAIKRIKIQONGGLAVEGHNGTLLAGAYTTSKLTITKLDKLNSEKLEKIEENAKKS 301
DB 265 A-KAIGKRIKNDVSLDNEADHNGSLISGAYILSNLTITKISAIKDSGELKAEIEKAKKS 323
OY 302 EEDTKLGEHQAQIGENTYDENAKKAILITDAKDKGAELKLEKRAVENLSKAAKEM 361
DB 324 EEDTKLGEHQAQIGENTYDENAKKAILITNDKTKGADELEKLEFESVKLSKAAKEM 383
OY 362 ANSVKEL 368
DB 384 TNSVKEL 390
```

```
RESULT 8
US-09-596-746a-72
; Sequence 72, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596.746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140.042
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 393
; TYPE: PRN
; ORGANISM: ospc Chimera
US-09-596-746a-72
```

Query Match 86.5%; Score 1554.5; DB 19; Length 393;
Best Local Similarity 87.2%; Pred. No. 5.5e-106;
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```
OY 2 ACNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDELATKAI 61
DB 26 SCNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDELATKAI 85
OY 62 GKRIKGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNELEKIDITAKOCSTEFNTKL 121
DB 86 GKRIKGNNGLEANOSKNTSLISGAYAISDLIAEKLNVLKNELEKIDITAKOCSTEFNTKL 145
OY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKLEKRAVENLSKAAODTLKNAVKE 181
DB 146 KSEHAVLGIDNLTDDNAORAILKHKANKDGAELKLEKRAVENLSKAAODTLKNAVKE 205
OY 182 TSPVHGNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDEL 241
DB 206 TSPVHGNNSGKDNASANSADSVKGNLTETISKRITESNAVLAKEVEETLLASIDEL 265
OY 242 ATRAIKRIKIQONGGLAVEGHNGTLLAGAYTTSKLTITKLDKLNSEKLEKIEENAKKS 301
DB 266 A-KAIGKRIKNDVSLDNEADHNGSLISGAYILSNLTITKISAIKDSGELKAEIEKAKKS 324
OY 302 EEDTKLGEHQAQIGENTYDENAKKAILITDAKDKGAELKLEKRAVENLSKAAKEM 361
DB 325 EEDTKLGEHQAQIGENTYDENAKKAILITNDKTKGADELEKLEFESVKLSKAAKEM 384
OY 362 ANSVKEL 368
DB 385 TNSVKEL 391
```

```

RESULT 9
US-09-596-746-38
; Sequence 38, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-38

Query Match
Best Local Similarity 85.5%; Score 1538; DB 19; Length 368;
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

QY 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
DB 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
QY 61 ICKKIGNNGLEANOSKNTSLSGAVAYSIDIAEKLAVLNKNEELKEKIDTAKOCSTPTNK 120
DB 61 ICKKIGNNGLEANOSKNTSLSGAVAYSIDIAEKLAVLNKNEELKEKIDTAKOCSTPTNK 120
QY 121 LKSEHAVLGIDNLTDDNAORAILKHKHANKDGAELKLEKFAVENLSKAADTLKNAVKE 180
DB 121 LKSEHAVLGIDNLTDDNAORAILKHKHANKDGAELKLEKFAVENLSKAADTLKNAVKE 180
QY 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
DB 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
QY 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKLTOKLDGLKNSKLEKIEENAKKC 300
DB 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKLTOKLDGLKNSKLEKIEENAKKC 300
QY 301 SEDFTKLLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADTL 360
DB 301 SEDFTKLLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADTL 360
QY 361 LANSVKEL 368
DB 361 LANSVKEL 368
;
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

```

```

; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-38

Query Match
Best Local Similarity 85.5%; Score 1538; DB 19; Length 368;
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

QY 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
DB 1 MACNSGKDGMSANSADSVKGPMLTEISKITTESNAVAVLAVKEVETLLASIDELATKA 60
QY 61 ICKKIGNNGLEANOSKNTSLSGAVAYSIDIAEKLAVLNKNEELKEKIDTAKOCSTPTNK 120
DB 61 ICKKIGNNGLEANOSKNTSLSGAVAYSIDIAEKLAVLNKNEELKEKIDTAKOCSTPTNK 120
QY 121 LKSEHAVLGIDNLTDDNAORAILKHKHANKDGAELKLEKFAVENLSKAADTLKNAVKE 180
DB 121 LKSEHAVLGIDNLTDDNAORAILKHKHANKDGAELKLEKFAVENLSKAADTLKNAVKE 180
QY 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
DB 181 LTPSPIVHGNSRGDNTSANSADSVKGPMLTEISKITTESNAVAVLAVKEIETLLASIDE 240
QY 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKLTOKLDGLKNSKLEKIEENAKKC 300
DB 241 LATKAIGKKIQONGGLAVEGHNGTLGAGVYTSKLTOKLDGLKNSKLEKIEENAKKC 300
QY 301 SEDFTKLLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADTL 360
DB 301 SEDFTKLLGEHQAQIGIENYTDENAKKAILITPAKDGAELKLEKFAVENLSKAADTL 360
QY 361 LANSVKEL 368
DB 361 LANSVKEL 368
;
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuitzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 391
; TYPE: PRF
; ORGANISM: OspC Chimera
US-09-596-746-76

```

Query Match 85.1%; Score 1530; DB 19; Length 391;
Best Local Similarity 86.4%; Pred. No. 3.5e-104;
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

QY 2 ACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 61
DB 25 SCNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 84
QY 62 GKRTGNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCSTEEFTNKL 121
DB 85 GKRTGNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCSTEEFTNKL 144
QY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTLKNAVKEL 181
DB 145 KSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTLKNAVKEL 204
QY 182 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 241
DB 205 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 264
QY 242 ATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTOKLDGLKNSERKEIKENAKKS 301
DB 265 ATKAIGKTI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCST 322
QY 302 EDPFKLEGEHAOLGIENVTDENAKKAILITDAKDGAEELEKLFKAVENLSKAADTL 361
DB 323 TEFNTKLSKSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTL 382
QY 362 ANSVKEL 368
DB 383 KNAVKEL 389

RESULT 12
US-09-596-746a-76
; Sequence 76, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596.746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746a-76

Query Match 85.1%; Score 1530; DB 19; Length 392;
Best Local Similarity 86.4%; Pred. No. 3.5e-104;
Matches 317; Conservative 21; Mismatches 27; Indels 2; Gaps 2;

QY 2 ACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 61
DB 26 SCNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKAI 85
QY 62 GKRTGNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCSTEEFTNKL 121
DB 86 GKRTGNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCSTEEFTNKL 145
QY 122 KSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTLKNAVKEL 181

DB 146 KSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTLKNAVKEL 205
QY 182 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 241
DB 206 TSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 265
QY 242 ATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTOKLDGLKNSERKEIKENAKKS 301
DB 266 ATKAIGKTI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCST 323
QY 302 EDPFKLEGEHAOLGIENVTDENAKKAILITDAKDGAEELEKLFKAVENLSKAADTL 361
DB 324 TEFNTKLSKSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTL 383
QY 362 ANSVKEL 368
DB 384 KNAVKEL 390

RESULT 13
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhulzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596.746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ospc Chimera
US-09-596-746a-26

Query Match 82.8%; Score 1489; DB 19; Length 374;
Best Local Similarity 83.4%; Pred. No. 3.4e-101;
Matches 312; Conservative 22; Mismatches 34; Indels 6; Gaps 2;

QY 1 MACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKA 60
DB 1 MACNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDELATKA 60
QY 61 IGKRTI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCSTEEFTN 119
DB 61 IGKRTI-GNNGLEANSKNTSLSGAVAIISDLIAEKLNVKNELEKIDTAKOCSTEEFTN 120
QY 120 KLSKSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTLKNAVKEL 179
DB 121 KLSKSEHAVLGIDNLTDDNAORAILKHKANKDGAEELEKLFKAVENLSKAADTLKNAVKEL 180
QY 180 ELTSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 234
DB 181 ELTSPYHGNNSGKDNASANSADSVKGPMLTEISKRTESNAVLAKEVEETLLASIDEL 240
QY 235 LASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTOKLDGLKNSERKEIKEXI 294
DB 241 LASIDELATKAIGKTIQONGGLAVEAGHNGTLLAGAVTISKLTOKLDGLKNSERKEIKEXI 300
QY 295 ENAKKSEDPFKLEGEHAOLGIENVTDENAKKAILITDAKDGAEELEKLFKAVENLSKAADTL 354
DB 301 ENAKKSEDPFKLEGEHAOLGIENVTDENAKKAILITDAKDGAEELEKLFKAVENLSKAADTL 360

OY 355 KAKEMLANSVKEL 368
|
Db 361 KAKEMLANSVKEL 374

RESULT 14

US-09-596-746-30
; Sequence 30, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match 82.6%; Score 1485.5; DB 19; Length 377;
Best Local Similarity 82.5%; Pred. No. 6.2e-101;
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

OY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEVEETLLASIDELATKA 60
|
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEINKKITDSNAVLAKEVEALLSIDETIAAKA 60
OY 61 ICKKI-GNNGLEPNOSKNTSLSGAVAISDLIAEKLNVKNEBELKEIDTAKOCSTEFN 119
|
Db 61 ICKKHONNGDLENNHNSLGAFAISTLIKOKLDGLKNGLEKIDAAKCCSTFTN 120
OY 120 KIKSEHAVLGIDNLTDDNARAILKHKANKDKGAELKLFRAVENLSKRAADTLKNAV 179
|
Db 121 KIKKHTDGLKRGVTDADAKKAILKANGTKTKGAELKLFESVEVLSKAAKEMLANSVK 180
OY 180 ELTSPYVHG-----NNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 231
|
Db 181 ELTSPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 240
OY 232 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKITOKLDGLKSEK 291
|
Db 241 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKITOKLDGLKSEK 300
OY 292 EKIEAKKCSDEFTKKLEGEHAOLGIEENVTDENAKKAILITDAKDKGAELKLFKAVE 351
|
Db 301 EKIEAKKCSDEFTKKLEGEHAOLGIEENVTDENAKKAILITDAKDKGAELKLFKAVE 360
OY 352 NLAKAKEMLANSVKEL 368
|
Db 361 NLAKAKEMLANSVKEL 377

RESULT 15

US-09-596-746a-30
; Sequence 30, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuizen, Daniel

; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-30

Query Match 82.6%; Score 1485.5; DB 19; Length 377;
Best Local Similarity 82.5%; Pred. No. 6.2e-101;
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

OY 1 MACNNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEVEETLLASIDELATKA 60
|
Db 1 MACNNSGKDGNTSANSADSVKGPMLTEINKKITDSNAVLAKEVEALLSIDETIAAKA 60
OY 61 ICKKI-GNNGLEPNOSKNTSLSGAVAISDLIAEKLNVKNEBELKEIDTAKOCSTEFN 119
|
Db 61 ICKKHONNGDLENNHNSLGAFAISTLIKOKLDGLKNGLEKIDAAKCCSTFTN 120
OY 120 KIKSEHAVLGIDNLTDDNARAILKHKANKDKGAELKLFRAVENLSKRAADTLKNAV 179
|
Db 121 KIKKHTDGLKRGVTDADAKKAILKANGTKTKGAELKLFESVEVLSKAAKEMLANSVK 180
OY 180 ELTSPYVHG-----NNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 231
|
Db 181 ELTSPVVAESPCKPSMVNNSGKDGNTSANSADSVKGPMLTEISKITSSNAVLAKEI 240
OY 232 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKITOKLDGLKSEK 291
|
Db 241 ETLASIDELATKAIGKKTIOQNGGLAVEGHNGTLLAGAVTTSKITOKLDGLKSEK 300
OY 292 EKIEAKKCSDEFTKKLEGEHAOLGIEENVTDENAKKAILITDAKDKGAELKLFKAVE 351
|
Db 301 EKIEAKKCSDEFTKKLEGEHAOLGIEENVTDENAKKAILITDAKDKGAELKLFKAVE 360
OY 352 NLAKAKEMLANSVKEL 368
|
Db 361 NLAKAKEMLANSVKEL 377

Search completed: March 18, 2002, 10:08:46
Job time: 974 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:35 ; Search time 55.5 Seconds

(without alignments)
149,211 Million cell updates/sec

Title: US-09-596-746A-42

Perfect score: 1798
Sequence: 1 MACNNSGKDGDNASNSADES.....AVENTLAKAKEMLANSVKEL 368Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	699	38.9	588	4	US-08-235-836C-122 Sequence 122, App
2	698.5	38.8	466	4	US-08-235-836C-107 Sequence 107, App
3	665	37.0	466	4	US-08-235-836C-110 Sequence 110, App
4	657.5	36.6	212	1	US-08-158-353-2 Sequence 2, Appl
5	644.5	35.8	209	4	US-09-196-293-15 Sequence 15, Appl
6	644.5	35.8	210	1	US-08-158-353-3 Sequence 3, Appl
7	639.5	35.6	210	4	US-08-235-836C-30 Sequence 30, Appl
8	637.5	35.5	210	4	US-08-209-603E-15 Sequence 15, Appl
9	600.5	33.4	212	1	US-08-031-295-2 Sequence 2, Appl
10	600.5	33.4	212	4	US-07-903-580-2 Sequence 2, Appl
11	588.5	32.7	212	1	US-08-158-353-4 Sequence 4, Appl
12	583.5	32.5	212	4	US-08-235-836C-34 Sequence 34, Appl
13	578.5	32.2	212	4	US-09-196-293-11 Sequence 11, Appl
14	578.5	32.2	212	4	US-08-209-603E-11 Sequence 11, Appl
15	563	31.3	207	4	US-08-235-836C-36 Sequence 36, Appl
16	553	30.8	209	4	US-08-235-836C-32 Sequence 32, Appl
17	168.5	9.4	1196	2	US-08-144-121-4 Sequence 4, Appl
18	168.5	9.4	1196	2	US-08-735-893-4 Sequence 4, Appl
19	152.5	8.5	630	4	US-08-973-462-9 Sequence 9, Appl
20	151.5	8.4	1181	1	US-08-053-614-4 Sequence 4, Appl
21	151.5	8.4	1181	1	US-08-316-397B-4 Sequence 4, Appl
22	151.5	8.4	1181	2	US-09-034-306-4 Sequence 4, Appl
23	151.5	8.4	1181	4	US-09-259-437-4 Sequence 4, Appl
24	151.5	8.4	1181	5	US-09-259-437-4 Sequence 4, Appl
25	151.5	8.4	2482	1	US-08-328-254-6 Sequence 6, Appl
26	150.5	8.4	3248	1	US-08-353-700-1 Sequence 1, Appl
27	150.5	8.4	3248	5	PCT-US95-16216-1 Sequence 1, Appl

28	146.5	8.1	1164	4	US-08-923-992A-2 Sequence 2, Appl
29	146.5	8.1	1786	4	US-08-973-462-8 Sequence 8, Appl
30	145	8.1	1147	3	US-08-470-260-5 Sequence 5, Appl
31	145	8.1	1147	3	US-08-471-491-5 Sequence 5, Appl
32	145	8.1	1147	4	US-08-466-662-5 Sequence 5, Appl
33	145	8.1	3289	2	US-08-477-451-2 Sequence 2, Appl
34	142	7.9	688	3	US-09-141-047-8 Sequence 8, Appl
35	141.5	7.9	3111	2	US-08-460-309-4 Sequence 4, Appl
36	141.5	7.9	3111	2	US-08-125-077-4 Sequence 4, Appl
37	141	7.8	2101	1	US-08-466-390-4 Sequence 4, Appl
38	141	7.8	2101	1	US-08-470-950-4 Sequence 4, Appl
39	141	7.8	2101	1	US-08-467-781-4 Sequence 4, Appl
40	141	7.8	2101	1	US-08-195-487-4 Sequence 4, Appl
41	141	7.8	2101	2	US-08-483-924-4 Sequence 4, Appl
42	141	7.8	2101	4	US-09-452-294-1 Sequence 1, Appl
43	141	7.8	2101	5	PCT-US93-06160-4 Sequence 4, Appl
44	138.5	7.7	1128	4	US-08-923-992A-6 Sequence 6, Appl
45	136.5	7.6	679	4	US-08-913-942-15 Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-122
Sequence 122, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Thereof
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ. ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-122

Query Match 38.9%; Score 699; DB 4; Length 588;
Best Local Similarity 41.5%; Pred. No. 5.1e-45;
Matches 195; Conservative 49; Mismatches 104; Indels 122; Gaps 17;
QY 7 GKQGNASNSADESVK-----PMLTETSKITTESNAVLAVER-----LLASID 54

Db 9 GSQKENDLNLEDSKSKSHQNAKODLPAYTEDSVSLFNGKIKFVSKKSNSSGKYDLRATID 68
Qy 55 ELATAIKKIGNNG--LEANO-SKNTSLSGAIAISDLIAEIKNLVLEKNELEKIDTA 110
Db 69 OVELGTGSDK--NMSGTLEGSKPKSKVLTIVSADLNTVLTLEAPDA-SNOKISSKV-TK 124
Qy 111 KQCS-TEFT--NKL-----KSEHVLGLDNLTD-DNAORAI----- 142
Db 125 KQSTTEETLRANKLDSKKLRPSNGTITLEYSOITPADNATRAVELKNSIKLEGLVYCK 184
Qy 143 -----LKKHAKND----- 150
Db 185 TTVEIKEGTVLKRERIEKDGKVFPLNDTAGSNKKTGKWEDESTPLTISADSKTKFDLVF 244
Qy 151 -----KGAELKLEKFAVENISKAQDPLKNAVKELTPIYHGNSSGDGNTS 198
Db 245 LTDGTTVOQYNTAGTSLEGSASEIKNLSE-----LKNALK--GHPM--GNSSGDGNTS 295
Qy 199 ANSADESVKGNPLTEISKRTESNAVLAVKEIETLLASIDELATRAIKGKIQQNGGLAV 258
Db 296 ANSADESVKGNPLTEISKRTESNAVLAVKEIETLLASIDELATRAIKGKIQQNGGLAV 355
Qy 239 EAGHGTLLAGATYTSKLTQKLDGLKNSKIEKIEKAKKSEDPYTKLEGEHQAQIGIE 318
Db 356 EYNHNSGLAGRYAISTLIKOKLDGLKN-EGLEKIDAKKCEFTYTKLEKHTDLGKE 414
Qy 319 NTDENAKKALITDAKOKGAELKLEKFAVENISKAQDPLKNAVKELTPIYHGNSSGDGNTS 368
Db 415 CVTDADAKKALITNGTITKGAELKLEFESVEVLSKAKKEMLANSVKEL 464

RESULT 2
US-08-235-836C-107
Sequence 107 Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESS: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235.836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-235-836C-107

Query Match 38.8%; Score 698.5; DB 4; Length 466;
Best Local Similarity 47.7%; Pred. No. 4,1e-45;
Matches 177; Conservative 43; Mismatches 114; Indels 37; Gaps 8;

Qy 19 ESVKGNPLTEISKRTESNAVLAVKEIETLLASIDELATRAIKGKIQQNGGLAV 77
Db 100 EYFKEDGTLVSKKVTSDKSSSTEKEFNEKEGVS-EKITRADGRLEGTGKSGSSKA 158
Qy 78 TSLSGAAYAISDLIAEIKNLVLEKNE--ELKEKIDTAKOCSEFTYTKLESEHVLGLDNL 134
Db 159 KEVLKGYVLEGTLLFAEKTTLVKEGTVLSKNSKSGEVSVELNDTSS--AATKRTAAM 216
Qy 135 DDNAORALKKHAKNDK-----GALEKLEKFAVENISKAQDPLKNA 177
Db 217 NDSTSTLTISADSKRTKQLVPLTDGTTVOQYNTAGTSLEGSASEIKNLSE-----LKN 271
Qy 178 VKELTSPYVHGNSSGDGNTSANSADSVKGNPLTEISKRTESNAVLAVKEIETLLAS 237
Db 272 LK-----MANNSSGDGNTSANSADSVKGNPLTEISKRTESNAVLAVKEIETLLAS 324
Qy 238 IDELATRAIKGKIQQNGGLAVAGHGTLLAGATYTSKLTQKLDGLKNSKIEKIEKAKK 297
Db 325 IDELATRAIKGKIQQNGGLAVAGHGTLLAGATYTSKLTQKLDGLKNSKIEKIEKAKK 383
Qy 298 KKCSDEFTYTKLEGEHQAQIGIENVDENAKKALITDAKOKGAELKLEKFAVENISKAQ 357
Db 384 KKCSDEFTYTKLEGEHQAQIGIENVDENAKKALITDAKOKGAELKLEKFAVENISKAQ 443
Qy 358 KEMLANSVKEL 368
Db 444 KEMLANSVKEL 454

RESULT 3
US-08-235-836C-110
Sequence 110 Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESS: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235.836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110

Query Match 37.0%; Score 665; DB 4; Length 466;
Best Local Similarity 46.9%; Pred. No. 1.3e-42;
Matches 165; Conservative 47; Mismatches 104; Indels 36; Gaps 8;

QY 1 MACNSGKGNASANSADSVKGNPLTEISKTTESNAVAVLAKVEYETLLASIDELATKA 60
DB 17 ISCNNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKVEYETLLASIDELATKA 76
QY 61 IGRKI-GNNGLEANOSKNTSLSGAVYASDLIAEKLNVLEKLEKIDTAKOCSTFTN 119
DB 77 IGRKIQNNGLDPEYNNHNSLLAGRYAISTLIKOKLDGLKNEGLKEKIDAKKCEFTFN 136
QY 120 KLRSEAVIGLNDLTDNAORALIKKHKANKGAAELKFRVAVENLSKAAODTLKNAK 179
DB 137 KLRKAKHTDGRKGVTDADAKKALKTNGTKGAEELKLFESVEVLSKAKEMLNLSYK 196
QY 180 ELTSPVHGNSGKDGNTSA---NSADSVKGNPLTEISKTTESNAVAVLAKVEETLL 235
DB 197 ELTSPVAVESPKPKQNVSSLDKNSVVDLPKMKVLSKKNKDGKY-----DLI 248
QY 236 ASIDELATRAIGKRIQNGSL-AVEAGHNGTLA-----GAYTISKLT 278
DB 249 ATVDKRLKGTSDKNNKSGVLEGVKADKSKVKLTISDDLQGTTLVEYFKEDGKTLVSKVT 308
QY 279 QKIDGKNSKLEKLEKLEMAKCSDETKKLEGHQ-LGIENVTDENAKKAI 329
DB 309 SK-DKSTEEKENGE---VSEKITTRADGTRLEYTGKSGSKAKKAEVL 355

RESULT 4

US-08-158-353-2
Sequence 2, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
NUMBER OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-2

Query Match 36.6%; Score 657.5; DB 1; Length 212;
Best Local Similarity 72.9%; Pred. No. 1.7e-42;
Matches 140; Conservative 20; Mismatches 29; Indels 3; Gaps 2;

QY 1 MACNSGKGNASANSADSVKGNPLTEISKTTESNAVAVLAKVEYETLLASIDELATKA 60
DB 17 ISCNNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKVEYETLLASIDELATKA 76
QY 61 IGRKIQNNGLDPEYNNHNSLLAGRYAISTLIKOKLDGLKNEGLKEKIDAKKCEFTFN 119
DB 77 IGRKIQNNGLDPEYNNHNSLLAGRYAISTLIKOKLDGLKNEGLKEKIDAKKCEFTFN 136
QY 120 KLRSEAVIGLNDLTDNAORALIKKHKANKGAAELKFRVAVENLSKAAODTLKNAK 177
DB 137 KLRKAKHTDGRKGVTDADAKKALKTNGTKGAEELKLFESVEVLSKAKEMLNLSYK 196
QY 178 VKELTSPVHGN 189
DB 197 VKELTSPVAVEN 208

RESULT 5

US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Molz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: 738,001US2
CURRENT APPLICATION NUMBER: US/09/196,293
EARLIER FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 209
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match 35.8%; Score 644.5; DB 4; Length 209;
Best Local Similarity 75.6%; Pred. No. 1.6e-41;
Matches 136; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 189 NSNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKVEYETLLASIDELATKA 248
DB 20 NSNGKDGNTSANSADSVKGNPLTEISKTTESNAVAVLAKVEYETLLASIDELATKA 79
QY 249 KIQONGGLAVEAGHNGTLGAYTISKLTQKIDGKNSKLEKLEMAKCSDETKKL 308
DB 80 KIQONGGLDPEYNNHNSLLAGRYAISTLIKOKLDGLKNEGLKEKIDAKKCEFTFNKL 138

STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 19920625
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BEHT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-580-2

Query Match 33.4%; Score 600.5; DB 4; Length 212;
Best Local Similarity 70.2%; Pred. No. 3,4e-38; Indels 1; Gaps 1;
Matches 127; Conservative 20; Mismatches 33;

QY 189 NNSGKDGNT-SANSADSVKGNLTETSKITTESNAVVLAVEIEFTLASIDELATK 247
DB 20 NNSGKDGDSASTNPADSAKGNLTETSKITDSNAFVLAVEIEFTLASIDELATK 79
QY 248 KRIQONGGLAVAGNGITLACATITSKITLITOKLDLGNSEKLEIKENAKCSEDFTK 307
DB 80 KRIQONGGLAVAGNGITLACATITSKITLITOKLDLGNSEKLEIKENAKCSEDFTK 139
QY 308 LGEHAQGIENVTDENAKKAILITDAKDKGALEKLFRAVENLAKAKEMANSYKE 367
DB 140 LAGHAQGIENVTDENAKKAILITDAKDKGALEKLFRAVENLAKAKEMANSYKE 199
QY 368 L 368
DB 200 L 200

RESULT 11
US-08-158-353-4
Sequence 4, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match 32.7%; Score 588.5; DB 1; Length 212;
Best Local Similarity 60.8%; Pred. No. 2,7e-37; Indels 3; Gaps 3;
Matches 130; Conservative 20; Mismatches 36;

QY 1 MACNNSGKDG-NASANSDESVKGNLTETSKITTESNAVVLAVEIEFTLASIDELATK 59
DB 17 ISCNNSGKDGDSASTNPADSAKGNLTETSKITDSNAFVLAVEIEFTLASIDELATK 76
QY 60 AIGKKI-GNGLNANOSKNTSLSGAVASDILAEKLVNKN-BELKEKIDTAKCSTEF 117
DB 77 AIGKIDNNNNLALNONGSLAGAVASITLITEKLSKLNELTELAKAKKCEEF 136
QY 118 TNKLSHVAVIDLNDNDNQRALIKKHNKDKGALEKLFRAVENLSKAADITKNA 177
DB 137 TNKLSGHADGKQDADTDHAKKAILKTHATTDKGAKEFDLFEVBGLKAAQVALTNS 196
QY 178 VKELTSPV 186
DB 197 VKELTSPV 205

RESULT 12
US-08-235-836C-34
Sequence 34, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-34

Query Match 32.5%: Score 583.5; DB 4; Length 212;
Best Local Similarity 68.3%: Pred. No. 6.5e-37;
Matches 129; Conservative 21; Mismatches 36; Indels 3; Gaps 3;

QY 1 MACNSGKDG-NASANSADSVKGNLTETSKITESNAVLAKEVETLLASIDELATK 59
DB 17 ICSNSGKGDASSTNPADSAKGPRLTEISKITDSNMFVLAVKEVETLVLSIDELAKK 76
QY 60 AIGKTI-GNNGLEANOSKNTSLSGAVYASDLAEKLVNKN-EELKEKIDTAKOCSTER 117
DB 77 AIGQIDNNNGIALNNGNSLAGAVYSTLTETKLSKLNLEELKTEITAKAKCSEEF 136
QY 118 TTKLSEHAVLGLDNLTDNAORAILKHKANKDGAEELEKLFKAVENSKAODTLKNA 177
DB 137 TTKLSGHADLGDKQDATHAKAAILKTHATTDKAKERKDFESVEGLKKAQVALTNS 196
QY 178 VKELTSPV 186
DB 197 VKELTSPV 205

RESULT 13
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: burgdorferi
CURRENT FILING DATE: 1998-11-19
EARLIER FILING DATE: 1994-03-10
EARLIER FILING DATE: 1992-06-19
EARLIER FILING DATE: 1990-12-21
EARLIER FILING DATE: 1989-12-22
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match 32.2%: Score 578.5; DB 4; Length 212;
Best Local Similarity 68.3%: Pred. No. 1.5e-36;
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

QY 1 MACNSGKDG-NASANSADSVKGNLTETSKITESNAVLAKEVETLLASIDELATK 59
DB 17 ICSNSGKGVGLISTNPADSAKGPRLTEISKITDSNMFVLAVKEVETLVLSIDELAKK 76
QY 60 AIGKTI-GNNGLEANOSKNTSLSGAVYASDLAEKLVNKN-EELKEKIDTAKOCSTER 117
DB 77 AIGQIDNNNGIALNNGNSLAGAVYSTLTETKLSKLNLEELKTEITAKAKCSEEF 136
QY 118 TTKLSEHAVLGLDNLTDNAORAILKHKANKDGAEELEKLFKAVENSKAODTLKNA 177
DB 137 TTKLSGHADLGDKQDATHAKAAILKTHATTDKAKERKDFESVEGLKKAQVALTNS 196
QY 178 VKELTSPV 186
DB 197 VKELTSPV 205

RESULT 14
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELLIA BURGDOFFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209, 603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862, 535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 212
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:

Query Match	31.3%;	Score 563;	DB 4;	Length 207;
Best Local Similarity	62.6%;	Pred. No. 2.2e-35;		
Matches 117;	Conservative 34;	Mismatches 32;	Indels 4;	Gaps 3;

	Matches	117	Conservative	34;	Mismatches	32;	Indels	4;	Gaps	3;
y	1	MACNNGKDGANSA	NDESVKGNLTETIS	KRTITENAVVLAVKEVETLLASIDELATKA	60	: : : : :	: : : : :	: : : : :	: : : : :	:
b	17	ISCNNGG-GDSASTP	DDESAKGPNTVTYS	KRTITDENAFLAVKEVEALLSIDET-SKA	73	: : : : :	: : : : :	: : : : :	: : : : :	:
y	61	IGKKIINN-LEANOSKN	TSLSGAVIDLAEARK	VNVKNEIKETIDTAOCSEFPN	119	: : : : :	: : : : :	: : : : :	: : : : :	:

[illegible]

THIS PAGE BLANK (USPTO)

S70287

outer surface protein C - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000

C:Accession: S70287

R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.

Mol. Microbiol. 18, 257-269, 1995

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease

A:Reference number: S70255; MUID:96296448

A:Accession: S70287

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Cross-references: EMBL:U42895; NID:9858723; PIDN:AB37003.1; PID:91695220

A:Experimental source: Strain 28354

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 48.7%; Score 876; DB 2; Length 193;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 NNSGKDNSTANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGK 248

DB 2 NNSGKDNSTANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGK 61

QY 249 KIQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKL 308

DB 62 KIQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKL 121

QY 309 EGEHAQGIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 368

DB 122 EGEHAQGIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 181

RESULT 3

I40153

outer surface protein C - Lyme disease spirochete (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000

C:Accession: I40153

R:Funf, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.

Infect. Immun. 62, 3213-3221, 1994

A:Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in I

A:Reference number: I40153; MUID:94314437

A:Accession: I40153

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-191 <RES>

A:Cross-references: EMBL:U08284; NID:9469561; PIDN:AAA21460.1; PID:9469562

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 48.0%; Score 863; DB 2; Length 191;

Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 190 NSGKGNSTANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGK 249

DB 1 NSGKGNSTANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGK 60

QY 250 IQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKLE 309

DB 61 IQONGGLAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKLE 120

QY 310 GEHAQGIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 368

DB 121 GEHAQGIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 179

RESULT 4

I40125

outer surface protein C - Lyme disease spirochete (strain M01) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000

C:Accession: I40125; S54195

R:Thiesen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

J. Bacteriol. 177, 3036-3044, 1995

A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A:Reference number: I40104; MUID:95286481

A:Accession: I40125

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178 <RES>

A:Cross-references: EMBL:X84779; NID:9793817; PIDN:CAA59250.1; PID:9793818

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 45.9%; Score 826; DB 2; Length 178;

Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 196 NTSANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGKIQONGG 255

DB 1 NTSANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGKIQONGG 60

QY 256 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKLEGEHAOL 315

DB 61 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKLEGEHAOL 120

QY 316 GIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 368

DB 121 GIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 173

RESULT 5

I40104

outer surface protein C - Lyme disease spirochete (strain 272) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000

C:Accession: I40104; S54184

R:Thiesen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

J. Bacteriol. 177, 3036-3044, 1995

A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A:Reference number: I40104; MUID:95286481

A:Accession: I40104

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178 <RES>

A:Cross-references: EMBL:X84785; NID:9793787; PIDN:CAA59256.1; PID:9793788

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 45.8%; Score 823; DB 2; Length 178;

Matches 170; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 196 NTSANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGKIQONGG 255

DB 1 NTSANSADSVKGPMLTEISKTTESNAVVLAVKEITLLASIDELATKAIGKIQONGG 60

QY 256 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKLEGEHAOL 315

DB 61 LAVEAGHNGTLLAGAYTISKLTOKLDGLKNSKLEKTEENAKKCSDEFTKKLEGEHAOL 120

QY 316 GIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 368

DB 121 GIENVTDENAKKAILITDAKDKGAELKFAVENLAKAKEMLANSVKEL 173

RESULT 6

I40124
outer surface protein C - Lyme disease spirochete (strain KIPP) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40124; S54194
R:Thielsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; MUID:95286481
A:Accession: I40124
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL:X84782; NID:9793815; PIDN:CAA59253.1; PID:9793816
C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 41.1%; Score 739; DB 2; Length 178;
Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Q 196 NTSANSADESVKGPMLTEISKRTESNAVLA VKEIETLLASIDELATRAIGKTIQONGC 255
|||||
Db 1 NTSANSADESVKGPMLTEISKRTESNAVLA VKEIETLLASIDELATRAIGKTIQONGC 60
Q 236 LAVEAGHNGTLLAGAVTTSKLTQKLDGLKNSKLEKLENNKCSDEFTKLEGEHAOL 315
|||||
Db 61 LAVEAGHNGTSLAGAVTTSKLTQKIRWENSEKITGKLENNKCSDEFTKLEGEHAOL 120
Q 316 GLENTDENAKKAILITDAKAKGAAELEKLEKAVENLAKAKEMLANSVKEL 368
|||||
Db 121 GLENTDENAKKAILITDAKAKGAAELEKLEKAVENLAKAKEMLANSVKEL 173

RESULT 7

S69927
outer surface protein C precursor - Lyme disease spirochete (strain PKa)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
A:Variety: Strain PKa
C:Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C:Accession: S69927; S72669
R:Jauris-Helpe, S.; Liegl, G.; Praec-Muric, V.; Roessler, D.; Schwab, E.; Soutschek, E.
J. Clin. Microbiol. 33, 1860-1866, 1995
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A:Reference number: I40047; MUID:95395018
A:Accession: S69927
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-210 <JAN>
A:Cross-references: EMBL:X69589
A:Experimental source: strain PKa
R:Jauris, S.
submitted to the EMBL Data Library, February 1994
A:Reference number: S72669
A:Accession: S72669
A:Molecule type: DNA
A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAN>
A:Cross-references: EMBL:X69589
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 36.1%; Score 648.5; DB 2; Length 210;
Best Local Similarity 76.1%; Pred. No. 3.2e-24;
Matches 137; Conservative 14; Mismatches 28; Indels 1; Gaps 1;

Q 189 NNSGDKNTSANSADSVKGPMLTEISKRTESNAVLA VKEIETLLASIDELATRAIGK 248
|||||
Db 20 NNSGDKNTSANSADSVKGPMLTEISKRTESNAVLA VKEIETLLASIDELATRAIGK 79
Q 249 KIQNGGLAVEAGHNGTLLAGAVTTSKLTQKLDGLKNSKLEKLENNKCSDEFTKEL 308

|||||
Db 80 KIHQNGGLDTENNHSGLAGAVTTSKLTQKLDGLKN-EGLEKTEAKKCSFTTNKL 138
Q 309 EGEHAOLGIENTYDENAKKAILITDAKAKGAAELEKLEKAVENLAKAKEMLANSVKEL 368
: : |||
Db 139 KKHNDLCKEGVTDADAKKAILITNTKTKRGADELCKLEFESVLSKAKEMLANSVKEL 198

RESULT 8

I40145
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40145
R:Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A:Title: Expression and sequence of outer surface protein C among North American isol
A:Reference number: I40143; MUID:95154673
A:Accession: I40145
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: EMBL:U04282; NID:92314881; PIDN:AAC4540.1; PID:9434666
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 35.9%; Score 646; DB 2; Length 211;
Matches 136; Conservative 18; Mismatches 32; Indels 2; Gaps 2;

Q 1 MACNNSGKDGK-ASANSADSVKGPMLTEISKRTESNAVLA VKEIETLLASIDELATRK 59
:::|||||
Db 17 ISCNNSGKDGKNAATPDADESVKGPMLTEISKRTESNAVLA VKEIETLLASIDELATRK 76
Q 60 AIGKRI-GNNGLEANOSKRTSLSGAYATSDIAELKLVNKKDELKRIIDTKKOSTEFT 118
|||||
Db 77 AIGKRIHQNNGLDTEENNHSGLAGAVTTSKLTQKLDGLKNELEKLEIAAVKCSSEFT 136
Q 119 NKLSSEHAYLGDINTDNOAQRALKKHNKKGAAELEKLEKAVENLSKAADTLKNAV 178
|||||
Db 137 NKLSSEHAYLGDINTDNOAQRALKKHNKKGAAELEKLEKAVENLSKAADTLKNAV 196
Q 179 KELTSPIV 186
|||||
Db 197 KELTSPIV 204

RESULT 9

G70218
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C:Accession: G70218; I40269; S37726; S70281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kevlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: AV0100; MUID:98065943
A:Accession: G70218
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <KIP>
A:Cross-references: GB:AE000792; NID:93253098; PIDN:AAC66329.1; PID:92689901; TIGR:BB
A:Experimental source: strain B31
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: I40269; MUID:96025162
A:Accession: I40269

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:D49497; NID:q707092; PIDN:BA08457.1; PID:q769684
R:Jauris-Heijpe, S.; Fuchs, R.; Motz, M.; Preac-Murisc, V.; Schwab, E.; Soutschek, E.; W
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)
A:Reference number: S37726; MUID:93268136
A:Accession: S37726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <JAU>
A:Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA9306.1; PID:g311392
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A:Reference number: S70255; MUID:96296448
A:Accession: S70281
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210 <LIV>
A:Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB3695.1; PID:g1695212
A:Experimental source: strain ip2
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.8%; Score 644.5; DB 2; Length 210;
Best Local Similarity 75.6%; Pred. No. 4.9e-24;
Matches 136; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 189 NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEIEITLASIDELATKAIGK 248
DB 20 NNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEIEITLASIDELATKAIGK 79

QY 249 KIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKEKTEIENAKKCESEFTKKL 308
DB 80 KIHQNGGLDTEHNNHNSILAGAYTISKLTITOKLDGKLN-EGIKETIDAKKCESEFTKKL 138

QY 309 EGEHAOLGIENVTDENAKKAILITDAAKGAELKLFKAVENTLAKAKEMLANSVKE 368
DB 139 KEKHTDLGEGVTDADAKKAILITKTGKAELKLFESVEVLAKAKEMLANSVKEL 198

RESULT 10
S70279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-193 <LIV>
A:Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
A:Experimental source: strain 25015
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.7%; Score 642; DB 2; Length 193;
Best Local Similarity 72.6%; Pred. No. 5.8e-24;
Matches 135; Conservative 16; Mismatches 33; Indels 2; Gaps 2;

QY 3 CNNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEIEITLASIDELATKA 61
DB 1 CNNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEIEITLASIDELATKA 60

QY 62 GKKT-CNNCLEANOSKNTSLSGAVALSDLAETKLVNKEELKEKIDTAKCSTETPNK 120
DB 61 KIKHNNGLDTEHNNHNSILAGAYTISKLTITOKLDGKNSKEKTEIENAKKCESEFTKK 120

QY 121 LKSEHAVGLDMLTDPNNQORATLKKHANKGKAELKLFKAVENTLAKAKEMLANSVKE 180
DB 121 LKSSHTLELKGQADDDAKKAILRTHNTKDGKAELKLFKAVENTLAKAKEMLANSVKE 180

QY 181 LTPSPV 186
DB 181 LTPSPV 186

RESULT 11
S70280
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: S70280
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; MUID:96296448
A:Accession: S70280
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-193 <LIV>
A:Cross-references: EMBL:L42868; NID:g858735; PIDN:AAB37011.1; PID:g1695226
A:Experimental source: strain Z57
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 35.6%; Score 640; DB 2; Length 193;
Best Local Similarity 76.2%; Pred. No. 7.2e-24;
Matches 138; Conservative 16; Mismatches 25; Indels 2; Gaps 2;

QY 189 NNSGKDGNTSANSADSVKGPMLTEISKRTTESNAVLAVKEIEITLASIDELATKAIGK 248
DB 2 NNSGKDGNTSANSADSVKGPMLTEISKRTTDSNAVLAVKEIEITLASIDELATKAIGK 60

QY 249 KIQONGGLAVEAGHNGTLLAGAYTISKLTITOKLDGKNSKEKTEIENAKKCESEFTKKL 308
DB 61 KIKNDGSLGDEAHNHNSILAGAYTISKLTITOKLDGKNSKEKTEIENAKKCESEFTKKL 120

QY 309 EGEHAOLGIENVTDENAKKAILITDAA-KDKGAELKLFKAVENTLAKAKEMLANSVKE 367
DB 121 KDNHAOLGIQGVTDENAKKAILITDAAKGAELKLFESVEVLAKAKEMLANSVKE 180

QY 368 L 368
DB 181 L 181

RESULT 12
I40144
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40144; S70282
R:Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A:Title: Expression and sequence of outer surface protein C among North American isol
A:Reference number: I40143; MUID:95154673
A:Accession: I40144
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664
R:Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995

Db 20 NNSGKDGNTSANSADSVKGPNLTEISKITTDNSNAVLAVKEVEVLSSIDELAKKAIK 79
QY 249 KIOONGSLAVEAGHNGTLLAGAYTISKLIYOKLDGLNSEKLEKIEINAKKCSSEFTK 308
Db 80 KIDONNALGTLDNHNHNSLGLAGAYTISALITEKLSIKDSELKAEIEKAKKCSSEFTK 139
QY 309 EGEHAOLGIENVTDENKKAAILITDAKDKGALELEKLFKAVENTLAKAKEMLANSVK 368
Db 140 SDNOAELGIENATDDNKAAILKTHNAKDKGALELVKLSSEVAGLLKAAOAILANSVK 199

Search completed: March 18, 2002, 09:56:49
Job time: 372 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:28 ; Search time 81.02 Seconds

(without alignments)
290.158 Million cell updates/sec

Title: US-09-596-746A-42

Sequence: 1 MACNNSCKDGNASANSADSE.....AVENLAKAKEMLANSVKEL 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA.New:*
1: /cgn2_6/ptodata/2/paa/BCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US05_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	640	35.6	211	6	US-09-974-992-5
2	588.5	32.7	212	6	US-09-974-992-7
3	567	31.5	207	6	US-09-974-992-3
4	176	9.8	837	6	US-09-815-242-5883
5	176	9.8	837	6	US-10-072-851-5883
6	176	9.8	837	6	US-09-815-242-13080
7	176	9.8	837	6	US-10-072-851-13080
8	170	9.5	1304	6	US-09-708-427-15045
9	170	9.5	1313	6	US-09-708-427-15044
10	169.5	9.4	2434	6	US-09-815-242-5835
11	169.5	9.4	2434	6	US-10-072-851-5835
12	169.5	9.4	6281	6	US-09-815-242-12996
13	169.5	9.4	6281	6	US-10-072-851-12996
14	168.5	9.4	1765	5	US-10-037-182-8
15	168.5	9.4	1765	5	US-09-938-275-6
16	168.5	9.4	1786	6	US-09-873-676-113
17	168.5	9.4	1786	6	US-09-561-709B-9
18	168.5	9.4	1786	6	US-10-037-182-6
19	167.5	9.3	1013	6	US-09-614-150-9159
20	163.5	9.1	1881	7	US-10-032-585-7646
21	163.5	9.1	1881	7	US-10-072-851-15590
22	162	9.0	1014	6	US-09-708-427-19883
23	162	9.0	1018	6	US-09-708-427-19882
24	162	9.0	1269	6	US-09-708-427-19881
25	161.5	9.0	730	6	US-09-708-427-3731

26	161.5	9.0	805	6	US-09-708-427-3730	Sequence 3730, Ap
27	161.5	9.0	841	6	US-09-708-427-3729	Sequence 3729, Ap
28	158.5	8.8	5795	7	US-09-815-242-12610	Sequence 12610, A
29	158.5	8.8	5795	7	US-10-072-851-12610	Sequence 12610, A
30	155.5	8.6	2056	6	US-09-614-150-4824	Sequence 4824, Ap
31	155	8.6	1038	6	US-09-708-427-32624	Sequence 32624, A
32	155	8.6	1042	6	US-09-708-427-32623	Sequence 32623, A
33	155	8.6	1043	6	US-09-708-427-32622	Sequence 32622, A
34	155	8.6	2076	6	US-09-815-242-5815	Sequence 5815, Ap
35	155	8.6	2076	7	US-10-072-851-5815	Sequence 5815, Ap
36	155	8.6	2186	6	US-09-815-242-12913	Sequence 12913, A
37	155	8.6	2186	7	US-10-072-851-12913	Sequence 12913, A
38	155	8.6	2437	6	US-09-815-242-5834	Sequence 5834, Ap
39	155	8.6	2437	7	US-10-072-851-5834	Sequence 5834, Ap
40	154.5	8.6	1144	6	US-09-708-427-15046	Sequence 15046, A
41	153	8.5	1725	5	US-10-037-182-12	Sequence 12, Appl
42	153	8.5	1786	5	US-09-938-275-7	Sequence 7, Appl
43	153	8.5	1786	5	US-10-037-182-10	Sequence 10, Appl
44	152.5	8.5	630	6	US-09-742-096-5	Sequence 5, Appl
45	152	8.5	1128	7	US-10-032-585-7627	Sequence 7627, Ap

ALIGNMENTS

RESULT 1
US-09-974-992-5
; Sequence 5, Application US/0974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Thelssen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel Ospc-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974, 992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180, 089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 5
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match 35.6%; Score 640; DB 6; Length 211;
Best Local Similarity 76.2%; Pred. No. 8.2e-37;
Matches 18; Conservative 16; Mismatches 25; Indels 2; Gaps 2;
QY 189 NNSGKQNTSANSADSEVKNPNTLEISKRTTESNAVLAVKEITETLLASIDELATKRG 248
DB 20 NNSGKQNTSANSADSEVKNPNTLEISKRTTESNAVLAVKEITETLLASIDELATKRG 78
QY 249 KIOONGLAVEACHNGTLLAGAYTISKRLYQKLDGKNSKLEKTIENARKCEDFTKRL 308
DB 79 KINQDSIGDEANHNESLLAGAYTISLTLQKLSKLSGKLEKTIENARKCEDFTKRL 138
QY 309 EGEHAQIGINVTDENAKKAILTTDA-KKGALELEKRAVENLAKAKEMLANSVKRE 367
DB 139 KDHAQDLGIGVTDENAKKAILANAGKDKVEELEKLSGSLSSKAKEMLANSVKRE 198
QY 368 L 368
DB 199 L 199
RESULT 2
US-09-974-992-7

```
Sequence 7, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
TITLE OF INVENTION: Novel OspC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia afzelii
US-09-974-992-7
```

```
Query Match 32.7%; Score 588.5; DB 6; Length 212;
Best Local Similarity 68.8%; Pred. No. 2.7e-33;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;
```

```
QY 1 MACNSGKDG-NASANSADSVGPNLTETISKRITESNAVLAVKEVETLLASIDELAK 59
Db 17 ISCNNSGKDGSDASTNPDESAGKPNLTETISKRITDSNAFLAVKEVETLLASIDELAK 76
QY 60 AIGKAT-GNNGLEANSKNTSLSGAVASDLAEKLVNKN-EELKEKIDTAKOCSTEF 117
Db 77 AIGKIDNNNGLAALNNGSILAGAYASTLITERKSLKMLEKTIKIAAKKCSF 136
QY 118 TTKKSEHAVLGIDNTDNOARAILTKRANKDKGAELKFKAVENISKAADTLKNA 177
Db 137 TTKKSGHADIKQDATTDHAKAAILKTHATTDKGAKKEKDFESVEGILKAQVALTNS 196
QY 178 VKELTSPV 186
Db 197 VKELTSPV 205
```

```
RESULT 3
Sequence 3, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
TITLE OF INVENTION: Novel OspC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 207
TYPE: PRT
ORGANISM: Borrelia garinii
US-09-974-992-3
```

```
Query Match 31.5%; Score 567; DB 6; Length 207;
Best Local Similarity 63.1%; Pred. No. 7.8e-32;
Matches 118; Conservative 33; Mismatches 32; Indels 4; Gaps 3;
```

```
QY 1 MACNSGKDGNSANSADSVKGNLTETISKRITESNAVLAVKEVETLLASIDELAK 60
Db 17 ISCNNSG--GDSASTNPDESAGKPNLTETISKRITDSNAFLAVKEVETLLASIDELAK 73
QY 61 ICKKIGNNG-LEANSKNTSLSGAVASDLAEKLVNKN-EELKEKIDTAKOCSTEF 119
Db 74 ICKKIKNDGTLEARNKNSILAGAYASTLITERKSLKMLEKTIKIAAKKCSF 133
QY 120 KUKSEHAVLGIDNTDNOARAILTKRANKDKGAELKFKAVENISKAADTLKNAVK 179
Db 134 KLDKSHAEIGSQYDDNNAKKAAILKTHGTRKDGAKAELELFRSLSKAQAALITNSVK 193
QY 180 ELTSPV 186
Db 194 ELTNPV 200
```

```
RESULT 4
US-09-815-242-5883
Sequence 5883, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5883
LENGTH: 837
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5883
```

```
Query Match 9.8%; Score 176; DB 6; Length 837;
Best Local Similarity 23.1%; Pred. No. 0.00026;
Matches 93; Conservative 77; Mismatches 154; Indels 78; Gaps 20;
```

```
QY 4 NNSGKDGNSANSADSVKGNLTETISKRITESNAV-VLAKE-VETLLASIDELAK 53
Db 100 NNLLEAVAKONANRAIDGLTSLNGPQAKLKEOVGQATTLPNVOTVYDNDATLNAKGLR 159
QY 54 DELATAIGKIGKIGNNGLEANSKNTSLSGAVASDLAEKLVNKN-EELKEKIDTAKOCSTEF 109
Db 160 DSIANEAT-IKAGQNTYDASQNKOTYNSAVTAARAILIGOTSPSMNAQEIQAQDV-T 217
QY 110 AKOCSEFTNKKL-----SEHAVLGIDNTDNOARAILTKRANKDKGAELKFKAVE 164
Db 218 AKQQAALNGEENLKTQATNAKQHLNGLSLDLD--AOKDAVKROI--EGATHVEVETQAN 272
```

```

Oy      4  NNSKDCDGNASNSADE---SVKSGNLTLEISKITTESNAV---VLAVKE-VETTLASI-----53
Db      100 NNNLEIVAKQNNNTAIDGLTSLNGCKQAKKLEKQVGAATTLRNQVQVRBNAQTLMTAMKGLR 159
Oy      54 DELATTKAIGKKTGNGGLEANOSKTTSLSCGAYAISDLAE-----KLAVLKNEELKEKIDT 109
Db      160 DSIANEAT-ITAGONYTDAOSKOTDYNSAVTAAKAIIGOTTSPTSMAOEINAKKDDV-T 217
Oy      110 AKOCSTEEFTNKLK-----SEHAUVGLDNLTDYDNNQRAITLKKHANKDKGAELEKFKRAVE 164
Db      218 AKQOALNGQEWLPTAOTFNAKOHNLGSLDLT--AQKRAVVRQI---EGATHVNNVVTQAO 272
Oy      165 NLS--KAAODTLKNAVVELTSPRIYHGNSSGCDGTSANSAKDESKPGNLTLEISKITTESN 222
Db      273 NADALNTAMTLKNGKID-----QWIKQGV-VNTDDADE-----AKRNNAYTN 313
Oy      223 AVVLAVAEIETLLASIDELATTKAIGKKTIOQNGGLAVEAG-----HNGTLLAGAY 271

```

[illegible]

Db 351 AVTQA-----EOLTKAQGPNTSKDG---VETALENVQRAKELNGNQNANAK 396
 QY 272 TISKLTITOKLDGKNSER--LKEKIEENAK--CSEDFTKKEGHAOLGIEN-VTDENA 325
 Db 397 TTKAKNALNLTISINNAKQKALKSQIEGATTVAGVQVSTSELTMTAMSNLQNGINDEEA 456
 QY 326 KKAIIITDAKDKGALELEKLEKAVENLAKAKEMLANSVKE 367
 Db 457 TKAALNGTONLEKAKQAHNTAIDGLSHLTNQAQKALQVQ 498

RESULT 7

US-10-072-851-13080
 ; Sequence 13080, Application US/10072851
 ; GENERAL INFORMATION:
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Xu, H. Howard
 ; APPLICANT: Foulkes, J. Gordon
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Roemer, Terry
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Boone, Charles
 ; APPLICANT: Bussey, Howard
 ; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
 ; FILE REFERENCE: ELITRA.028A
 ; CURRENT APPLICATION NUMBER: US/10/072,851
 ; CURRENT FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: 60/257,636
 ; PRIORITY FILING DATE: 2001-02-09
 ; NUMBER OF SEQ ID NOS: 15811
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13080
 ; LENGTH: 875
 ; TYPE: PRP
 ; ORGANISM: Staphylococcus aureus
 ; US-10-072-851-13080

Query Match 9.8%; Score 176; DB 7; Length 875;
 Best Local Similarity 23.1%; Pred. No. 0.00028;

Matches 93; Conservative 77; Mismatches 154; Indels 78; Gaps 20;

QY 4 NNSGKDGASANSADSE--SVKGNLTETISKITESNAV--VLAVKE-VETLLASI----- 53
 Db 137 NNNLEVAQNANNTAIDGLTSLNGPQAKLKQVGAATLTPVQTVPRDQAOTLNTAMGLR 196
 QY 54 DELTAKIGKIGNGNGLEANSKNTSLSGAVATSDLAIE---KLVTKMELEKEDT 109
 Db 197 DSIINENF--IKAGONYDASQNKQTDYNSAVTAKAIIIGTTSQNSMAQETNOAKDOV--T 254
 QY 110 AKQOSTEFTNKLK-----SEHAVLGLDNLTDNAPRALIKKHANKDKGALELEKFAVE 164
 Db 255 AKQALNGQELRTAQTNAKQHLNGLSPLLD--AQKDAVKRQI---EGATVNEVTQAQN 309
 QY 165 NLS--KKAQDTLKAQVVELTSPYIVHGNSGKDGNTSANSADSESVKGNLTETISKITESN 222
 Db 310 NADALNTAMNLTNGKIDP-----ONTIKQG--VNEFTDADE-----AKRNAVTN 350
 QY 223 AVVLAIVEIETILASIDELATKAIGKIIQONGGLAVEAG-----HNGTLGAY 271
 Db 351 AVTQA-----EOLTKAQGPNTSKDG---VETALENVQRAKELNGNQNANAK 396
 QY 272 TISKLTITOKLDGKNSER--LKEKIEENAK--CSEDFTKKEGHAOLGIEN-VTDENA 325
 Db 397 TTKAKNALNLTISINNAKQKALKSQIEGATTVAGVQVSTSELTMTAMSNLQNGINDEEA 456

QY 326 KKAIIITDAKDKGALELEKLEKAVENLAKAKEMLANSVKE 367
 Db 457 TKAALNGTONLEKAKQAHNTAIDGLSHLTNQAQKALQVQ 498

RESULT 8

US-09-708-427-15045
 ; Sequence 15045, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. Alexandrov et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09
 ; NUMBER OF SEQ ID NOS: 85364
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15045
 ; LENGTH: 1304
 ; TYPE: PRP
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc-feature
 ; LOCATION: 1..1304
 ; OTHER INFORMATION: Xaa is any amino acid
 ; NAME/KEY: misc-feature
 ; LOCATION: 1..1304
 ; OTHER INFORMATION: Ceres Seq. ID 1828628
 ; US-09-708-427-15045

Query Match 9.5%; Score 170; DB 6; Length 1304;
 Best Local Similarity 22.8%; Pred. No. 0.0012;

Matches 91; Conservative 80; Mismatches 136; Indels 92; Gaps 17;

QY 4 NNSGKDGASANSADSE-----VKGNLTETISKITESNAVLAVK-----EYET 48
 Db 85 NEKLEKALAAQRAEESFEVEKFRAVELEQAGLEAVQKDVTSKNEDESIRSHALDISA 144
 QY 49 LLASIDELATFAIGKIGNGNGLEANSKNTSL--LSGAVATSDLAIEKLVTKME----- 101
 Db 145 LLSITTEL-----QRYKHEISMTADAKNKLASHAEETKTAETLHAETLASELRLK 198
 QY 102 -----ELKEKIDTAKOSTEFTNKLKSEHAVLGLDNLTDNAPRALIKKHANKDKGALE 155
 Db 199 ALLGSKKEKALE-----GNETVSKIKSE-----IELLGLLEKVSILLESILKQEBGLVE 248
 QY 156 LEKL-----FRAVENLSKRAQDTLKNAYKELTSPYIVHGNSGKDGNTSANSADSEVKG--P 209
 Db 249 QLKVDLEAKKAESCNTNSVSEWENKRYHELEKEV-----EESNRSKSSASEMESVMK 301
 QY 210 NLTEISKITESNAVLAIVE--IETLLASIDELAT--KAIGKIIQONGGLAVEAGHNGTL 266
 Db 302 QLAELNHYLETTKDANAQKEKIELLEKTEIAQRTDLEEGROV----- 345
 QY 267 LAGAYTISKLTITOKLDGKNSERLKEKIEENAKKCSSEDFTKKEGHAOL--GIENVTDENA 325
 Db 346 -----CIAKEBASKELEU--VESIKSELSISQ---EKTALDNEKAKATSNIQULDDQRT 395
 QY 326 KKAIIITDAKDKGALELEKLEKAVENLAKAKEMLANS 364
 Db 396 ELSTEL-----ERCKYVEERSKDMESLTALOEASTES 429

RESULT 9

US-09-708-427-15044
 ; Sequence 15044, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. Alexandrov et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427

```

: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 15044
: LENGTH: 1313
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 1..1313
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc.feature
: LOCATION: 1..1313
: OTHER INFORMATION: Ceres Seq. ID 1828627
: US-09-708-427-15044

```

```

Query Match          9.5%; Score 170; DB 6; Length 1313;
Best Local Similarity 22.8%; Pred. No. 0.0012;
Matches 91; Conservative 80; Mismatches 136; Indels 92; Gaps 17;

```

```

4 NNSGKDGNASANSADSE-----VKGPNLTEISKITTESNAVAVLAK-----EVEI 48
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 NKKLAKALAAOKRAEESFEVEKFRAVELBOAGLEAVOKKDVTSKNELESIRSOHALDISA 153
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49 LASIDELATKAIGKKTGNNGLEANSKNTSL--LSGATAISDLIAEKLVAKNE----- 101
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
154 LISTTELE-----ORVKHELMTADAKKALSHAEAKIKAEIHAEKELLIASELGRUK 207
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
102 -----ELKEKIDTAKOCSTFTNKLKSEHAVGLDNLTDNAQRALILKHKAKGAAE 155
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
208 ALLGSEKEEAIE-----GNEIYSKLAKSE-----IELNGELEKVSILLESLEKOGELYE 237
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 LEKL-----EKAVENLSKAADPTLKNVAKELTSPIVHGNSGKDGNTSANSADSEVKG--P 209
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
258 QLKVDLEAKMAECSNTSSVEEEMKNVHELEKEV-----EESNRKSASSESMSVVK 310
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
210 NLTETSKTTESNAVAVLAKKE--IETLASIDELAT--KAIGKTIQONGGLAVAGANGTL 266
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
311 QALHLHVHETKSDNAOKKEKTELEKTIEAORTDLEEGRV----- 354
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
267 LAGAVTISKLTIOKLDGLKNSKLEKIEENAKKSEDFTKLGEHAOL--GIENVTDENA 325
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 -----CIAKEEASKLENL--VESIKSELEISQ---EKKRALDNEKRAAISNIONLDORT 404
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
326 KKAALITDAKDKGAELKLEKFAVENLAKAKEMLAN 364
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 ELSTEL-----ERCKVEEESKKDMESLTLAQEASTES 438
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

US-09-815-242-5835
: Sequence 5835, Application US/09815242
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Travick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26

```

```

: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5835
: LENGTH: 2434
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
: US-09-815-242-5835

```

```

Query Match          9.4%; Score 169.5; DB 6; Length 2434;
Best Local Similarity 23.9%; Pred. No. 0.0029;
Matches 94; Conservative 76; Mismatches 144; Indels 79; Gaps 21;

```

```

4 NNSGKDGNASANSADSE-----SVKGNLTEISKITTESNAV--VLAKVE-VETLLAST----- 53
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
702 NNNLEVAQKQANATIDGLTSLNGPQAKLKEQYGAATTLPPNQTVDNMQTLNTAKGLR 761
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 DELATATGKKTGNNGLEANSKNTSLSGATAISDLIAE-----KINVLKNEELKEKIDT 109
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
762 DSTANET--IKAQONTYDASONKOTDYNSAVTAKAKIIGTTSPSMAQETINAKDOV--T 819
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 AKOCSTFTNKLK-----SEHAVGLDNLTDNAQRALILKHKAKGAALEKFAVE 164
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
820 AKQALNGOENLFTADTNKQHLNGSLDLD--AQNDVAVRQI---EGATHVAVENVAQN 874
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
165 NLS--KAADPTLKNVAKELTSPIVHGNSGKDGNTSANSADSEVKGPNLTEISKTTESN 222
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
875 NADALMTAMTNLKNGIOD-----QNTIKQC--VNETDADE-----AKRNAYTN 915
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
223 AVVLAKKEIETLLASIDELATKAIGKKTQONGGLAVEAG-----HNTGLAGAY 271
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
916 AVTQA-----EQILNKAOGPNTSKDG--VETALENVORAKNEINGNONVANAK 961
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
272 TISKLTIOKLDGLKNSK--LKEKIEENAKK--CSEDFTKLGEHAOLGIEN--VTDENA 325
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
962 TTKKNNLNLSTINNAOKELKQIEGATTVAGVNOYSTASSELNTAMSNLQNGINDEAA 1021
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
326 KKAALITDAKDKGAELKLEKFAVENLAKAA 357
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1022 TKAAQKYTDADREKQTAYNDAVTAAKTLLDKTA 1054
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
US-10-072-851-5835
: Sequence 5835, Application US/10072851
: GENERAL INFORMATION:
: APPLICANT: Carr, Grant J.
: APPLICANT: Xu, H. Howard
: APPLICANT: Foulkes, J. Gordon
: APPLICANT: Zamudio, Carlos
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Travick, John D.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Roemer, Terry
: APPLICANT: Jiang, Bo
: APPLICANT: Boone, Charles
: APPLICANT: Bussey, Howard
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
: FILE REFERENCE: ELITRA.028A
: CURRENT APPLICATION NUMBER: US/10/072,851
: CURRENT FILING DATE: 2002-02-08

```

PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5835
LENGTH: 2434
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-072-851-5835

Query Match 9.4%; Score 169.5; DB 7; Length 2434;
Best Local Similarity 23.9%; Pred. No. 0.0029;
Matches 94; Conservative 76; Mismatches 144; Indels 79; Gaps 21;

QY 4 NNSGDNASANSAD---SVKPNLTESKITESNAV-VLAKE-VETLLASI----- 53
DB 702 NINLEVARONATPAIDGLTSLNGPQAKLKEQVGQATLTPNVQTVRDNAQTINTAMKGLR 761
54 DELATKAIGKIGKNGGLEANSKNTSLSGAVALDLAE-----KLNVKNELEKIDT 109
DB 762 DSIAHEAT-IRAGONTYDASOKKOTDYNASVTAAKAIIGOTTSPSMNOETNOAKDQV-T 819
QY 110 AKCSTFTNKLK-----SEHAVLGIDNLTDMNQRAILKKHANKDKGALEKLEKAVE 164
DB 820 AKQALNGOENLRTAQTNAKQHLNGLSLDLD--AQKDAVKRQI---EGATHVNEVTOAQN 874
QY 165 NLS--KAADTLKNAVKELTSPIVHNGNSGDKNTSANSADSVKPNLTESKITESN 222
DB 875 NADALNTMTNKNQIDQ-----QNTIKQG-VNFTDAE-----AKNNAYTN 915
QY 223 AVILAWEIETLLASIDELATKAIGKIQOONGLAVEAG-----HNGTLLAGAY 271
DB 916 AVTQA-----EOLINKAQGPNTSKDS---VETALENVQRAKNELNQNVANAK 961
QY 272 TISKILITOKLGLKNSER--LKEKTENAK--CSEDTFKLGEHAOLGIEN-VTDENA 325
DB 962 TTKAKNALNNLTSINNAQKEALKSQIEGATTVAGVNOVSTASELNTAMSNLNGINDEAA 1021
QY 326 KKAJ-LITDAKDKGALEKLEKFAVENLAKAA 357
DB 1022 TKAOKYTDADREKOTAYNDAVTAAKTLLDKTA 1054

RESULT 12
US-09-815-242-12996
Sequence 12996, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zysek, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12996
LENGTH: 6281
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match 9.4%; Score 169.5; DB 6; Length 6281;
Best Local Similarity 23.9%; Pred. No. 0.0099;
Matches 94; Conservative 76; Mismatches 144; Indels 79; Gaps 21;

QY 4 NNSGDNASANSAD---SVKPNLTESKITESNAV-VLAKE-VETLLASI----- 53
DB 1481 NINLEVARONATPAIDGLTSLNGPQAKLKEQVGQATLTPNVQTVRDNAQTINTAMKGLR 1540
54 DELATKAIGKIGKNGGLEANSKNTSLSGAVALDLAE-----KLNVKNELEKIDT 109
DB 1541 DSIAHEAT-IRAGONTYDASOKKOTDYNASVTAAKAIIGOTTSPSMNOETNOAKDQV-T 1598
QY 110 AKCSTFTNKLK-----SEHAVLGIDNLTDMNQRAILKKHANKDKGALEKLEKAVE 164
DB 1599 AKQALNGOENLRTAQTNAKQHLNGLSLDLD--AQKDAVKRQI---EGATHVNEVTOAQN 1653
QY 165 NLS--KAADTLKNAVKELTSPIVHNGNSGDKNTSANSADSVKPNLTESKITESN 222
DB 1654 NADALNTMTNKNQIDQ-----QNTIKQG-VNFTDAE-----AKNNAYTN 1694
QY 223 AVILAWEIETLLASIDELATKAIGKIQOONGLAVEAG-----HNGTLLAGAY 271
DB 1695 AVTQA-----EOLINKAQGPNTSKDS---VETALENVQRAKNELNQNVANAK 1740
QY 272 TISKILITOKLGLKNSER--LKEKTENAK--CSEDTFKLGEHAOLGIEN-VTDENA 325
DB 1741 TTKAKNALNNLTSINNAQKEALKSQIEGATTVAGVNOVSTASELNTAMSNLNGINDEAA 1800
QY 326 KKAJ-LITDAKDKGALEKLEKFAVENLAKAA 357
DB 1801 TKAOKYTDADREKOTAYNDAVTAAKTLLDKTA 1833

RESULT 13
US-10-072-851-12996
Sequence 12996, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zysek, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
FILE REFERENCE: ELITRA 028A
CURRENT APPLICATION NUMBER: US/10/072,851
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12996
LENGTH: 6281

TYPE: PRF
ORGANISM: Staphylococcus aureus
US-10-072-851-12996

Query Match 9.4%; Score 169.5; DB 7; Length 6281;
Best Local Similarity 23.9%; Pred. No. 0.0099;
Matches 94; Conservative 76; Mismatches 144; Indels 79; Gaps 21;

QY 4 NNSGDKGNSANSAD---SVKPNLITEISKITESNAV--VLAKV-VETLLAST-----53
DB 1481 NHLEFAKQNAVTALDGLSLNGPOKAKLEQVQATPLPNQVTDNMQTLTAAKGLR 1540
QY 54 DELATFAIKKIKGNNGLENOSKNTSLSGAVASLDLIE---KLVNLEKNEKIKDT 109
DB 1541 DSIFANFAT-IKAGQNTYDASQNKQTDYNSAVTAAKAIIGQTSPPSMAOEINQAKQV-T 1598
QY 110 AKQCSFEFTNKLK-----SEHAVLGDNLTDDNAQRAIKKHNKDKGALEKLFKAVE 164
DB 1599 AKQALNGENLRTAQTNAKQHLNGSLDLD--AQKDAVKRQI---EGATHVEVYQAQN 1653
QY 165 NLS--KAAODTLKNAVKELTSPYVHGNSGDKGNTSANSDESVKGNLITEISKITESN 222
DB 1654 NADALTAMTNKNGIQD-----QNTIKQG-VNFTDADE-----AKRRAVYTN 1694
QY 223 AVYLVAKLEIETLLASIDELATKAIGKIKQONGGLAVEAG-----HNGTLLAGAY 271
DB 1695 AVTQA-----EQILNKAQGNPNTSKD--VETALENVQARKNELNGQNVANAK 1740
QY 272 TISKLTQKLDGLKNSK--LKEKINAKK--CSEDFTKLEGEHAQGIEN-VTDENA 325
DB 1741 TTKAKNALNLTSTNNQAKALKSQIGCATTVAGVNOVSTTASLNTAAMSILONGINDEAN 1800
QY 326 KRAI-LITDAKDKGALEKLEKRAVENLAKAA 357
DB 1801 TKAQKTYTDADREKQYAVNDVATAKTLDDKTA 1833

RESULT 14
US-10-037-182-8
Sequence 8, Application US/10037182
GENERAL INFORMATION:
APPLICANT: Tryggvason, Karl
APPLICANT: Dol, Masayuki
TITLE OF INVENTION: Recombinant Laminin 10
FILE REFERENCE: 99-274-F
CURRENT APPLICATION NUMBER: US/10/037,182
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,449
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/279,282
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1765
TYPE: PRF
ORGANISM: Homo sapiens
US-10-037-182-8

Query Match 9.4%; Score 168.5; DB 7; Length 1765;
Best Local Similarity 24.6%; Pred. No. 0.0023;
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

QY 2 ACNNSGKDC-NASANSADSVKGNLITEI-----SKKITESNAVVLAVKREVEYLLAS 52
DB 1384 SCSETECGGPNCTDEGERKCGPGGLVTVAHNAHQAMDLDQVLSLALVEEOLSKM 1443
QY 53 IDELATKAIGKIKGNNGLENOSKNTSLSGAVASLDLIEKLVNLEKNEKIKDTAKQ 112
DB 1444 VSEAKLRAD-----EAKQSAEDILLK-----TNATREKMD-KSNEELRLMLIQIRN 1488

QY 113 CSTEFTNKLKSEHAVLG-----LDNLTDNNAQRA-----ILKHNKDKG 152
DB 1489 FLTQDSADLDSIEAVANEVLEKEMPESTPOQLQTLTDEREVESLSQVEYTLQHSADTA 1548
QY 153 AAELKLFKRAVENLSKAAODT-----LKNVAVELTSPYHGNSGDKGNTSANSADSV 206
DB 1549 RAEM--LLEBAKRAKSAVDVKTADNVKALEBAEKAAQVAARAKIK-----QADEDI 1599
QY 207 KGNP--LTEISKITESNAVVL-AVKEIETLLASIDELATKAIGKIKQONGGLAVEAGHN 263
DB 1600 QGTQNLITSTIESEPTAASEETLFNQSORISELERVVEELKRA-----AQNSG---BAEYI 1651
QY 264 GTLLAGAYTISKL---ITQKLDGLKNSKLEKIEIEN--AKKCEDFTKLEGEHAQGI 318
DB 1652 EKVY---YTVKQSAEDVKKTLQD-ELDEKVK-KVENLAKKITEESADARRKRAEMLONEAK 1706
QY 319 NVTDENAKKAILITDAAK-----DKGALEKLFKRAVENLAKKAKEMLA 362
DB 1707 TLLAQANSKLOLKLDEKREKEDNQRLEBK-AQELARLEGEVSLKIDISQKVA 1759

RESULT 15
US-09-938-275-6
Sequence 6, Application US/09938275
GENERAL INFORMATION:
APPLICANT: Gerardo Castillo
APPLICANT: Alan Snow
TITLE OF INVENTION: Therapeutic and Diagnostic Applications
TITLE OF INVENTION: The laminin and laminin-derived protein fragments
FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938,275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1786
TYPE: PRF
ORGANISM: Homo Sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Swissprot P07942
DATABASE ENTRY DATE: 1988-08-01
US-09-938-275-6

Query Match 9.4%; Score 168.5; DB 5; Length 1786;
Best Local Similarity 24.6%; Pred. No. 0.0023;
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

QY 2 ACNNSGKDC-NASANSADSVKGNLITEI-----SKKITESNAVVLAVKREVEYLLAS 52
DB 1405 SCSETECGGPNCTDEGERKCGPGGLVTVAHNAHQAMDLDQVLSLALVEEOLSKM 1464
QY 53 IDELATKAIGKIKGNNGLENOSKNTSLSGAVASLDLIEKLVNLEKNEKIKDTAKQ 112
DB 1465 VSEAKLRAD-----EAKQSAEDILLK-----TNATREKMD-KSNEELRLMLIQIRN 1509
QY 113 CSTEFTNKLKSEHAVLG-----LDNLTDNNAQRA-----ILKHNKDKG 152
DB 1510 FLTQDSADLDSIEAVANEVLEKEMPESTPOQLQTLTDEREVESLSQVEYTLQHSADTA 1569
QY 153 AAELKLFKRAVENLSKAAODT-----LKNVAVELTSPYHGNSGDKGNTSANSADSV 206
DB 1570 RAEM--LLEBAKRAKSAVDVKTADNVKALEBAEKAAQVAARAKIK-----QADEDI 1620
QY 207 KGNP--LTEISKITESNAVVL-AVKEIETLLASIDELATKAIGKIKQONGGLAVEAGHN 263
DB 1621 QGTQNLITSTIESEPTAASEETLFNQSORISELERVVEELKRA-----AQNSG---BAEYI 1672
QY 264 GTLLAGAYTISKL---ITQKLDGLKNSKLEKIEIEN--AKKCEDFTKLEGEHAQGI 318
DB 1673 EKVY---YTVKQSAEDVKKTLQD-ELDEKVK-KVENLAKKITEESADARRKRAEMLONEAK 1727

Tue Mar 19 10:57:56 2002

us-09-596-746a-42.rapn

Page 8

```

Oy      319  NVTDENAKKALLITDAK-----DKGAEELEKFEKAVENIAKAKAKEMLA 362
      : : | : | : | | | | : | : | : |
Db      1728 TLLAQANSKIQLDLDERKEYEDNORYLEDR-AQELAREGVEYRSLLIKDISQVTA 1780

```

Search completed: March 18, 2002, 09:58:33
Job time: 391 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:55 ; Search time 39.62 Seconds

(without alignments)
340.552 Million cell updates/sec

Title: US-09-596-746a-42

Perfect score: 1798

Sequence: 1 MACNNGKCGKGNASANSADES.....AVENLAKAKEMLANSVKEL 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	644.5	35.8	210	1	OSCL_BORBU
2	588.5	32.7	212	1	OSCL_BORBU
3	346	19.2	214	1	VM03_BORBU
4	325.5	18.1	215	1	VM03_BORBU
5	175	9.7	1790	1	US01_YEAST
6	173.5	9.6	1167	1	CAGA_HELPY
7	168.5	9.4	1167	1	LMB1_HUMAN
8	167.5	9.3	1013	1	A60D_DROME
9	161.5	9.0	944	1	NUP1_YEAST
10	158.5	8.8	1433	1	REST_CHICK
11	158	8.8	1182	1	CGA2_HELPY
12	157.5	8.8	483	1	M6_STRPY
13	157.5	8.8	1251	1	RBP2_PLAIV
14	157	8.7	1391	1	MST2_DROME
15	155.5	8.6	2017	1	MYSN_DROME
16	155	8.6	2022	1	ANT1_ONCVO
17	154.5	8.6	2349	1	TPR_HUMAN
18	153	8.5	1786	1	LMB1_MOUSE
19	152	8.5	1639	1	LMG1_DROME
20	151.5	8.4	1427	1	REST_HUMAN
21	151.5	8.4	1966	1	MYSB_CAEBL
22	151.5	8.4	3210	1	CENF_HUMAN
23	150	8.3	1509	1	MYSN_ACACA
24	149.5	8.3	1978	1	MYHB_CHICK
25	148.5	8.3	564	1	M12_STRPY
26	148.5	8.3	679	1	YKR9_YEAST
27	148	8.2	492	1	M5_STRPY
28	147	8.2	1186	1	SMC_BACSU
29	146.5	8.1	1164	1	BAG_STRAG
30	146.5	8.1	1189	1	SCIT_CHICK
31	146.5	8.1	1230	1	SMC3_YEAST
32	146.5	8.1	1957	1	Y086_SCHPO
33	145.5	8.1	705	1	YNP9_CAEBL

34	145	8.1	662	1	TEPA_BACSU
35	145	8.1	1147	1	CGAL_HELPY
36	145	8.1	1186	1	CAGA_HELPY
37	143.5	8.0	539	1	M21_STRPY
38	143.5	8.0	918	1	YMOB_CAEBL
39	143.5	8.0	1233	1	YF16_YEAST
40	143.5	8.0	1972	1	MYHB_HUMAN
41	143	8.0	724	1	HMHR_HUMAN
42	142.5	7.9	2116	1	MYS2_DICDI
43	142	7.9	1084	1	MYS2_RABIT
44	142	7.9	1526	1	MYS2_SCHPO
45	141.5	7.9	3110	1	LMA2_HUMAN

ALIGNMENTS

RESULT	ID	OSCL_BORBU	STANDARD:	PRT:	210 AA.
1	OSCL_BORBU	007337			
AC	OSCL_BORBU	15-DEC-1998 (Rel. 37, Created)			
DT	DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DR	DR	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	GN	OSPC OR BB19.			
OS	OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	OG	Plasmid Ip54.			
OC	OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.			
OX	OX	NCBI_TaxID=139;			
RN	RN	[1]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-ATCC 35210 / B31;			
RX	RX	MEDLINE-9339332; PubMed-809841;			
RA	RA	Jauris-Helpke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,			
RA	RA	Will G., Wilske B.;			
RT	RT	"Genetic heterogeneity of the genes coding for the outer surface			
RT	RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.";			
RL	RL	Med. Microbiol. Immunol. 182:37-50(1993).			
RN	RN	[2]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-ATCC 35210 / B31;			
RX	RX	MEDLINE-9339332; PubMed-8478108;			
RA	RA	Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,			
RA	RA	Schwab E., Wanner G.;			
RT	RT	"Immunological and molecular polymorphisms of OspC, an immunodominant			
RT	RT	major outer surface protein of Borrelia burgdorferi.";			
RL	RL	Infect. Immun. 61:2182-2191(1993).			
RN	RN	[3]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-ATCC 35210 / B31;			
RX	RX	MEDLINE-94041630; PubMed-8225587;			
RA	RA	Padilla S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;			
RT	RT	"Molecular characterization and expression of p23 (OspC) from a North			
RT	RT	American strain of Borrelia burgdorferi.";			
RL	RL	Infect. Immun. 61:5097-5105(1993).			
RN	RN	[4]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-ATCC 35210 / B31;			
RX	RX	MEDLINE-96025162; PubMed-7494039;			
RA	RA	Fukunaga M., Hamase A.;			
RT	RT	"Outer surface protein C gene sequence analysis of Borrelia			
RT	RT	burgdorferi sensu lato isolates from Japan.";			
RL	RL	J. Clin. Microbiol. 33:2415-2420(1995).			
RN	RN	[5]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-ATCC 35210 / B31;			
RX	RX	MEDLINE-98065943; PubMed-9403685;			
RA	RA	Frazer C.M., Castjens S., Huang W.M., Sutton G.G., Clayton R.A.;			
RA	RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,			
RA	RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,			
RA	RA	Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,			
RA	RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,			

RT are highly polymorphic." ;
 RL Mol. Microbiol. 6:3299-3311(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04786; AAA22964.1; -
 DR Interpro; IPR001800; Lipoprotein_6.
 DR Pfam; PF01441; Lipoprotein_6; 1.
 DR Prodom; PD001149; Lipoprotein_6; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Outer membrane; Lipoprotein; Signal; Plasmid.
 FT SIGNAL 1 18 PROBABLE.
 FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 214 AA: 22541 MW: 5158351024657C7 CRC64;
 Query Match 19.2%; Score 346; DB 1; Length 214;
 Best Local Similarity 42.1%; Pred. No. 1.2e-09;
 Matches 82; Conservative 40; Mismatches 57; Indels 16; Gaps 5;
 QY 1 MACNSGKDGNSANSADSESVKGP---NLTEISKKITESNAVLAVKEVEETLLASIDEL 56
 DB 17 MNCNNGPE-----LKSDEVAKSDGTVDLAKVSKIKKESAFASVKEVETLVSVDEL 71
 QY 57 ATKAKGKIKGNNGNGLANSKNTSLSGAYASDLAEKLVNLT-----NEELKEKIDRA 110
 DB 72 AKKAGKIKKNGDGLTERAGONGSLAGVSVAVKIKVGALETTSIGISNELKRTITEV 130
 QY 111 KOCSTEFNKLKSEHAVVLGLDNLTDNMRALIKKHANKKGALEKLFKAVENLSKAA 170
 DB 131 KSKAFAFLNKLKDGHTLGGKKSADSDPTKAIKKNOSDKTKASLEALNTAVDALTKAA 190
 QY 171 QDTLKNVAKELTSP 185
 DB 191 EGVEVAIKELTAPV 205
 QY 3 BORNE 4
 ID 3 BORNE STANDARD: PRT; 215 AA.
 AC Q02448;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
 GN VMP3.
 OS Borrelia hermsli.
 OC Plasmid.
 CC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 CC NCBI_TaxID=140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SSP. HSI SEROTYPE 3;
 RA MEDLINE=93133110; PubMed=1484486;
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
 RT "Subtelomeric expression regions of Borrelia hermsli linear plasmids
 RT are highly polymorphic." ;
 RL Mol. Microbiol. 6:3299-3311(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 CC FROM ONE SURFACE EXPOSED VMP TO ANOTHER.

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 CC ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP24.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04789; AAA22967.1; -
 DR Interpro; IPR001800; Lipoprotein_6.
 DR Pfam; PF01441; Lipoprotein_6; 1.
 DR Prodom; PD001149; Lipoprotein_6; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Outer membrane; Lipoprotein; Signal; Plasmid.
 FT SIGNAL 1 18 PROBABLE.
 FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 215 AA: 23139 MW: 684C74D35F87C771 CRC64;
 Query Match 18.1%; Score 325.5; DB 1; Length 215;
 Best Local Similarity 42.3%; Pred. No. 9.8e-09;
 Matches 80; Conservative 32; Mismatches 62; Indels 15; Gaps 5;
 QY 189 NNSGKDGNTSANSADSESVKGP---NLTEISKKITESNAVLAVKEVEETLLASIDELATK 244
 DB 20 NNGNPE-----LKSDEVAKSDGTVDLAKVSKIKKESAFASVKEVETLVSVDEL 73
 QY 245 AIGKKIQ-ONGGLAVENGHTLLAGAVTISKLTOKLGLKN-----SEKLKEIKENAKK 299
 DB 74 AIGKKIKKNDNSNEDNDHNGSLIAGVFYILTVKAKLTSLQIIGISDELKTEVGVK 133
 QY 300 CSDFPKKLEGEHAQCIENVTDENAKKALITDAKKDKGALEKLFKAVENLSKAA 359
 DB 134 ESFAFTYQVSKRTDLAKGVTDAHAKSALIVDGTGKGAABLIRLNTAIDELAKAAND 193
 QY 360 MLANSVKEL 368
 DB 194 AVETVVKEL 202
 QY 5
 ID USOL_YEAST STANDARD: PRT; 1790 AA.
 AC P25386;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.
 GN USOL OR INT1 OR YDL058W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RA MEDLINE=91185402; PubMed=2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 RA Yamasaki M.;
 RT "A cytoskeleton-related gene, usol, is required for intracellular
 RT protein transport in Saccharomyces cerevisiae." ;
 RL J. Cell Biol. 113:245-260(1991).
 RN [2]
 RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
 RA Kendrick K.E.;
 RT Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.

```

RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -1- SIMILARITY: BELONGS TO THE VDP/USOI/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X54378; CAA38253.1;
CC EMBL: L03188; AAB00143.1;
CC EMBL: U53668; AAB6659.1;
CC PIR: A38455; A38455.
CC HSSP: P80220; 1DIP.
CC SGD: S0002216; USOI.
CC Interpro: IPR002017; Spectrin.
CC Transport: Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
CC KW DOMAIN 1 724 GLOBULAR HEAD.
CC FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
CC FT DOMAIN 991 1790 CHARGED (HYPER-HYDROPHILIC).
CC FT DOMAIN 1172 1786 DISPENSABLE FOR THE PROTEIN FUNCTION.
CC FT CONFLICT 847 847 ASP/GLU-RICH (ACIDIC).
CC FT CONFLICT 924 924 G -> E (IN REF. 2).
CC FT CONFLICT 1253 1253 V -> K (IN REF. 2).
CC FT CONFLICT 1319 1319 I -> V (IN REF. 2).
CC FT CONFLICT 1461 1461 N -> S (IN REF. 2).
CC FT CONFLICT 1581 1581 G -> S (IN REF. 2).
CC FT CONFLICT 1600 1600 I -> V (IN REF. 2).
CC FT CONFLICT 1661 1661 R -> S (IN REF. 2).
CC FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
CC SEQUENCE 1790 AA; 206424 MW; 6CE2B216B9FD4818 CRC64;

```

```

Query Match 9.7%; Score 175; DB 1; Length 1790;
Best Local Similarity 22.7%; Pred. No. 0.51;
Matches 106; Conservative 66; Mismatches 158; Indels 136; Gaps 19;

```

```

OY 3 CNGSKDG-----NASANSDESVKGNLTETSKIT-----TESNAVYLVK 44
DB 929 CNGSKDG-----NASANSDESVKGNLTETSKIT-----TESNAVYLVK 986
OY 45 E-----VETLLASIDELATKAIGKIGNNGLEAN--QSKNTSLGAVYISDLIAEK 94
DB 967 ESKNESSIQLSNLQKIDMSQEKENFOIGSIEKNTIQLKTT-----ISLDEQTK 1038
OY 95 LNVLKNEE-----LKEKIDTAKOCSTEFNTKLS-----EHAVALGLDM- 133
DB 1039 EETISKSSKDEYESQISLKEKLETATANDENVNKISLTKTREELAEALAYKNLK 1098
OY 134 -----TDDNAORAT-----LKHANKDKG-----AAELEKTFKAVENUS 167
DB 1099 NELETETSEKALKEVENEHEHLEKEKIQLEKATEFTKQOLNSLRANLESLEKHEDLA 1158
OY 168 ---KRAODTLKNAVVELSPIVHGNGSKDGTNSANSADSVKGNP----- 210
DB 1159 AQLKYEEOIANKEKQYVEEISQLN-----DETTSQOENESIKKKKNDLEGEVYKAMKSTS 1214
OY 211 LTESKTTESNAVYLVAVEIE-----TLLASIDELATKAIGKIQOQNGCLAVEAGH 262

```

```

DB 1215 EEOSNLKSEIDALNIQIKELKKRNETNEASLSIESKSEVETV--KIKF---LQDECNF 1269
OY 263 NGTLLAGAYTISKIITQKLDGNKSEKLEKENAKKSCSEDFTKLGEHAQIGIENVTD 322
DB 1270 K-----EKEVSLDEKIKLASDEKKNKYLELQKESKINEEDAKTELKIQLEKI 1319
OY 323 ENAKKALITDAAKDKGAELKFKAVENILAKAKEMLANVEL 368
DB 1320 TNLSTK-----AKEKSELSRLKTSSEERKNAEOLERLKNKEI 1358

```

RESULT 6

```

CAGA_HELPUP STANDARD; PRT; 1167 AA.
ID CAGA_HELPUP
AC 09ZLTI;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN)
DE (CAG PATHOGENICITY ISLAND PROTEIN 26).
GN CAGA OR CAI OR CAG26 OR JHP0495.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Molr D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC OR FUNCTION OF THE CYTOTOXIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AEO01483; AAD06073.1;
CC KW Antigen; Complete proteome.
CC FT DOMAIN 246 249 POLY-THR.
CC FT DOMAIN 882 889 POLY-ASN.
CC SEQUENCE 1167 AA; 129729 MW; FD5E86B81CEBD0F2 CRC64;

```

```

Query Match 9.6%; Score 173.5; DB 1; Length 1167;
Best Local Similarity 22.28%; Pred. No. 0.36;
Matches 105; Conservative 81; Mismatches 140; Indels 147; Gaps 25;

```

```

OY 5 NSGKGNASANSADSVKGNLTETSKIT-----VETLLASID 54
DB 693 NINMDLKFSPKSPDFKNGKN-KDPSKAEFLTKALKSVKQDGINPEWISVENLMAALN 751
OY 55 ELATKAIGKIGNNGLEANOSKNTSLGAVYISDLIAEKLVNLEKKEKIDTAKO-- 112
DB 752 EF-----KNGKNKDFSKVTQ-----AKSDLENSIKDVIINQKIKTDVNDLQAV 795
OY 113 -----CSTEFNTKLSKSEHVALGLDITDQ--AORALIKHANKDKGAELEKFKAVEN-- 165
DB 796 SVAKATGDFSG---VEQALADLKNFSKQOLAQAQAKNDPFGKNKA---LYOSVKNQV 848
OY 166 -----LSKRAODTL-----KNAVVELSPIVHGNGSKDGTNSANSADSVKGNP 213
DB 849 NCTLVGNGLSKRAEATTLTKSNFSIDIKELNAKL--GNFNNNNNGLENGTE-----PIYTO 901

```


FT	DISULFID	886	895	BY SIMILARITY.
Query Match			9.4%;	Score 168.5; DB 1; Length 1786;
Best Local Similarity			24.6%;	Pred. No. 0.99;
Matches 102; Conservative			61;	Mismatches 160; Indels 91; Gaps 21;
QY	2	ACNNSGKDQ-NASANSADSVKGPNUETL-----SKKITESNAVYLAKVEYETLLAS	52	
DB	1405	SCSETECGGPNCTDGEKRCGPGCGGLVYVAHNAWOKRAMLDDVDYSLALAEVQSLSKM	1464	
QY	53	IDELATKAIGKIKGNNGLEANOSKNTSLSGAYASDLIAKTLNVLKNEELKEKIDTAQK	112	
DB	1465	VSEAKLRAD-----EAKOSADIEDLK-----TNAFKEMD-KSNEELRNLIKIDRN	1509	
QY	113	CSEEFNFKKSEHAVLG-----LDNLTDNAQA-----ILKIRANKDG	152	
DB	1510	FLTQDSADSDSLAEAVENLKMEMPSTPOQJONTLEDIREVESLSOVEYTLQHSADIA	1569	
QY	153	AAELEKLFRAVENLSKAQDT-----LKNVKELTSPYVGNNSGKDQNTSANSASDV	206	
DB	1570	RAEM-LLEAKRASKASDVYKVTADMYKALAELEKAKOVAEAKIR-----QADEDI	1620	
QY	207	KAPN--LPEISKITEESNAVYL-AKELETTLASIDELATKAIGKIQONGSLAVEAGHN	263	
DB	1621	OGTQNLTLTSEETVASEETLFNASQRISELEERVVEELKRKA-----AONSG--EAEYL	1672	
QY	264	GTLLAGAVYTIKSL--ITQKIDGLKNSKEKTEKEN--AKKCEDPTKKEGHAQIGIE	318	
DB	1673	EKVY--TYVQOSADVKKTLIDG-ELDEKYK-KVENLIAKTEESADARRAKEMLQNEAK	1727	
QY	319	NVTDENAKKALLIYDAK-----DKGALEKLFRAVENLAKAKEMLA	362	
DB	1728	TLLAANSKIQTLNDLEKRYEDNORYLEDK-AQELARLEGEVSRLKIDISQKVA	1780	
RESULT	8			
ID	A60D_DROME	STANDARD:	PRT:	1013 AA.
AC	P91927; Q9W160;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	CALCIUM-BINDING MITROCHONDRIAL PROTEIN ANON-60DA.			
GN	ANON-60DA OR CG4589.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
CC	Ephyridioidea; Drosophilidae; Drosophila.			
NC	NCBI_TaxID=7227;			
RP	[1]			
RP	SEQUENCE FROM N.Y.			
RC	STRAIN=BRKRELEY.			
RC	MEDLINE=20196006; PubMed=10731132;			
RX	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Mananlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson S.C., Miklos G.L.G.,			
RA	April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,			
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Duan A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Mayan-Hocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Giolek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Talbot M., Kalush E.,			

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Modarity C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sigenklamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RP [2]
 RN SEQUENCE OF 626-944 FROM N.A.
 RC TISSUE-Ovary;
 RC MEDLINE=99168769; PubMed=10071211;
 RA Caggese C., Ragone G., Perrini B., Moschetti R., de Pinto V.,
 RA Calzai R., Barsanti P.;
 RT "Identification of nuclear genes encoding mitochondrial proteins:
 RT isolation of a collection of *D. melanogaster* cDNAs homologous to
 RT sequences in the Human Gene Index database.";
 RL Mol. Gen. Genet. 261:64-70(1999).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS IN POSITIONS 920 AND 930.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE003464; AAF47217.1; -
 CC EMBL: Y10912; CA71853.1; ALT_FRAME.
 CC FlyBase: FBgn0019886; CG4598.
 DR InterPro: IPR001813; 60s_ribosomal.
 DR InterPro: IPR002024; Bacterioferritin.
 DR InterPro: IPR003871; DUF223.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000861; REM_repeat.
 DR InterPro: IPR002555; RFL.
 DR InterPro: IPR001236; 1dh.
 DR Pfam: PF00428; 60s_ribosomal; 1.
 DR Pfam: PF02721; Bacterioferritin.
 DR Pfam: PF02723; DUF223; 1.
 DR Pfam: PF00036; ehand; 2.
 DR Pfam: PF02185; HRI; 1.
 DR Pfam: PF00056; 1dh; 1.
 DR Pfam: PF01605; RFL; 1.
 DR SMART: SMO0054; EPH; 2.
 DR PROSITE: PS00018; EF_HAND; 2.
 KM Mitochondrion: Calcium-binding. Repeat.
 FT CA_BIND 700 711 EF_HAND 1 (POTENTIAL).
 FT CA_BIND 763 794 EF_HAND 2 (POTENTIAL).
 FT CONFLICT 687 687 E -> Q (IN REF. 2).
 FT CONFLICT 690 690 K -> G (IN REF. 2).
 FT CONFLICT 717 717 K -> H (IN REF. 2).
 FT CONFLICT 736 736 E -> D (IN REF. 2).
 FT CONFLICT 740 740 K -> R (IN REF. 2).
 FT CONFLICT 761 761 K -> E (IN REF. 2).
 FT CONFLICT 943 943 G -> S (IN REF. 2).
 SQ SEQUENCE 1013 AA: 113579 MW: 8C710380263F262E CRC64;

Query Match 9.3%; Score 167.5; DB 1; Length 1013;
 Best Local Similarity 22.4%; Pred. No. 0.6;
 Matches 77; Conservative 63; Mismatches 147; Indels 57; Gaps 11;
 QY 29 ISKRITSNVAVLAVKEVEITLASIDELATRAIKKIGNGNGLEANSQKNTSLSGAVAIS 88
 DB 609 VKPEVRSSRAAKLLVNNVNMISGLDVL-----NDLEARQHQIQOAESSDYAAS 658
 QY 89 DLIAEKLVKNEBELKIDTAQOCSTFPNKKLSHNAVJGLDNLTDNARAILKRAAN 148
 DB 659 SPVPEQOVHIDELVATIRMKKESASDERFVGVGD-----LVKLEAD 702
 QY 149 KDGALELEKLFRAVENLSAADDTLNAVKELT-----SPIYGNNSGKDN 196
 DB 703 KD-GVYSVNEITRAVOSIDREATYDKQLEETELLISKLASRRRHEIYHIDLMNNIK 761
 QY 197 TSANSADESQVGNPLIEISKITSNVAVLAVKEIETLASIDELATRAIKKIQOONGL 256
 DB 762 VLKETSDEA-RLKHIEAVLEKFPADKGVYTVNDIRRVLESIGRDNKLSDAIEE---- 816
 QY 257 AVERGHNGTLLAGATYTSKITOKLDGLKNSKIKERIEANAKKSEDPFKLSEH-AQL 315
 DB 817 LISLDEKQVLAQAEQKIEKAIAK---SMKEAEKLSKSEVDKADK---DLSKLVNDIHDSAK 870
 QY 316 GIENVTDENAKKAILITDAKDKGALELEKLFRAVENLAKAKE 359
 DB 871 EIQDIANEMRDKERTVDPKAKELKA---EPAFK---DTAKTLKD 908
 RESULT 9
 ID NUFL_YEAST STANDARD; PRT; 944 AA.
 AC P32380;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NUFL PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110).
 GN NUFL OR SPC110 OR YDR356W OR D9476.3.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RA MEDLINE=92176232; PubMed=1541631;
 RA Mirzayan C., Copeland C.S., Snyder M.;
 RT "The NUFL gene encodes an essential coiled-coil related protein that
 RT is a potential component of the yeast nucleoskeleton.";
 RL J. Cell Biol. 116:1319-1332(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94064779; PubMed=7503995;
 RA Kilmartin J.V., Dyos S.L., Kershaw D., Finch J.T.;
 RT "A spacer protein in the Saccharomyces cerevisiae spindle poly body
 RT whose transcript is cell cycle-regulated.";
 RL J. Cell Biol. 123:1175-1184(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favellio A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jeter M.,
 RA Johnson D., Johnson L., Langston Y., Latreille P., Le T.,
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifkin L., Riles L., Taich A., Trevisan E., Vignati D.,
 RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
 CC ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT
 CC IS ESSENTIAL FOR GROWTH.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
 CC NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE

CC NUCLEOLUS.
 CC -1- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: 211582; CAAT7668.1; -
 DR EMBL: X73287; CAAS1733.1; -
 DR EMBL: U28372; AAB64791.1; -
 DR PIR: S26710; S26710.
 DR PIR: S34288; S34288.
 DR SGD: S0002764; NUP1.
 DR Coiled coil; Nuclear protein; Phosphorylation.
 DR DOMAIN 164 791 COILED COIL.
 DR 54 59 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 DR DOMAIN 726 731 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 DR 742 747 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 DR DOMAIN 731 944 ARG/TYR-RICH.
 DR ARGININE
 DR SEQUENCE 944 AA; 111781 MW; 04FAA074BB8A0BC8 CRC64;

Query Match 9.0%; Score 161.5; DB 1; Length 944;
 Best Local Similarity 23.2%; Pred. No. 1;
 Matches 100; Conservative 71; Mismatches 167; Indels 93; Gaps 21;

DB 6 SKDDNANASNADESVKGNLFEISKITNESNAV-VLAVKEVETLASTIDE-LAKKAIGK 63
 DB 91 SFRSGN-VDKSRKRLIDLDKKDVPMSOPLEQREHOMKERRDRLESKLLGK 145
 DB 64 K-----IGNNGLEANSKNTSLSGAYASIDIAEKLAVLKN-EELKEKIDYAKOC 114
 DB 146 RHITYANDISNKELEYINEIKSLK-----HEIKELREKENDTLNVDLEETDCLKNRL 200
 DB 115 TEFTKIKSEHNAVGLDNLTDNAORALLKHKANDKAAEELKFAV-----ENLS 167
 DB 201 QALEKELDAKKNITV-SKRVDDHS--GCIEEOMERKIAELERKLVKQVLELENN 257
 DB 168 -----KAAODTLKNAVEKELTSPVHGNNSKDGNTS-----ANSAD-SVKGP 210
 DB 258 DVQSLKSKDEKLKMLNNE--LKSNAEKOTQLEFKKNEKLNELKIKIDE 315
 DB 211 L-TEISKITNESNAVLAKEITEL-----ASIDELATRAIGKIQO-----NGC 255
 DB 316 MDLQKOKONESKRLKDELNELETFSENGSOSAKENELKMLKIAELEEEISTKNSQ 375
 DB 256 LAVEAGHGTLLAGAVTISKILITQKLDL-----KNSKLEKIKENAKKSE----- 302
 DB 376 LIAREKGLASLAQUTQLESKLNDRDQSGREELKTKNTKLDINDIAEEIVSKDER 435
 DB 303 -DFTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGALE--KLFAVENILAKAA 357
 DB 436 IIDQKVKYKQLENDLFVTKTHSEK---TITD-----NELSKOKLILILENDLKVA 485
 DB 358 KEMLANSVKEL 368
 DB 486 QEKYSKMEKEL 496

RESULT 10
 REST_CHICK STANDARD; PRT; 1433 AA.
 AC 042184; 042228; 057563; 057564;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).
 GN RSN.
 OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus
 CC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98137792; PubMed=9469933;
 RA Griparic L., Volosky J.M., Keller T.C. III;
 RT "Cloning and expression of chicken CLIP-170 and restin isoforms."
 RN Gene 206:195-208(1998).
 RN [2]
 RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
 RC TISSUE=Pectoralis muscle;
 RA Griparic L., Keller T.C. III;
 RT "Identification and expression of two novel CLIP-170/Restin isoforms
 RT expressed predominantly in muscle."
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
 CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
 CC CYTOSKELETON (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF014012; AAC60344.1; -
 DR EMBL: AF020764; AAC60345.1; -
 DR EMBL: AF045650; AAC03547.1; -
 DR EMBL: AF045651; AAC03548.1; -
 DR InterPro: IPR000938; CAP-GLY.
 DR InterPro: IPR001878; ZnF_CCHC.
 DR Pfam: PF01302; CAP-GLY; 2.
 DR SMART: SM00343; ZnF_C2HC; 1.
 DR PROSITE: PS00845; CAP-GLY 1; 2.
 DR Cytokeleton; Microtubules; Coiled coil; Alternative splicing.
 FT DOMAIN 79 121
 FT 144 207 CAP-GLY 1.
 FT DOMAIN 235 277 SER-RICH.
 FT 305 332 CAP-GLY 2.
 FT 351 1353 SER-RICH.
 FT 1414 1427 COILED COIL (POTENTIAL).
 FT 458 492 CCHC-BOX.
 FT 458 492 MISSING (IN SHORT ISOFORM).
 FT 803 803 TÖRKLEHARKTEQSLFEKTRADKIQRELEDT ->
 FT 803 803 RKROISEDPEN (IN ISOFORM CLIP-170(11)).
 FT 458 458 S -> GSSSKVS (IN ISOFORM CLIP-170(11)).
 FT 458 458 T -> RKROISEDPEN (IN ISOFORM CLIP-
 FT 458 458 170(11+35)).
 FT 309 309 K -> R (IN REF. 2; AAC03547).
 FT 440 440 E -> V (IN REF. 2; AAC03548).
 FT 440 440
 SQ SEQUENCE 1433 AA; 161026 MW; 5631CEB683498E23 CRC64;

Query Match 8.8%; Score 158.5; DB 1; Length 1433;
 Best Local Similarity 22.5%; Pred. No. 2; 2; Indels 67; Gaps 16;
 Matches 86; Conservative 78; Mismatches 152;

DB 18 DESVKGPNLFEISKITNESNAVLAKEVETLLASIDELATKAIGKKTGNNGLEANSKN 77
 DB 508 DLALRVEVAELNGRLSSK-----HIDVDVTSLSLQETIS--SLQEMAAAGFE-HQREM 560
 DB 78 TSLISGAYASIDIAEKLAVL-----KNELEKIDTANQSTFTN--KLKSEHA 126
 DB 561 SLIKEKESSEALRKRIKTLASNERMGKENSLETKIDHANKNSDVLEKSKLSA 620
 DB 127 VLGLDNLTDNAORALLKHKANDKGA--AELEKLFAVENILSKAODTLKN----- 176

```

Db      621  IASHQAAEE-----LKVSFNKGVGAQOTAEFAELKTQMEKVKLDYENENSKLTKQENE 674
Oy      177  -----ANKELTSPDIVHGNNSGCKGNGNANSADSQVGPULTELSKITESMNAVLYAKE 230
Db      675  KSQHLKETELAKULLEVTEKEKOTLENLKAKLETSEVEDOHVHEMEDTLNKLQENELKYKE 734
Oy      231  IETLLASIDELATRAIGKIQONSGCLAVEAGHNGTL---LAGAYTISKLITQKLDGL 285
Db      735  LDVLOAKNE-OATLKLISLTQQ-----IRASEEKLLDIALAOKNANSGCKLEIKRLEO 788
Oy      286  NSEKLTKEKIEIAKKCSDEFKTLGEGHAQGIENVTDENAKKAILITDAKDAKGAEL-- 343
Db      789  AAKRQIQIULETEK--VSNLTRELOKGRKQKL-----LDLEKNLSAANYQVKDLSLEKQL 839
Oy      344  --EKLFRKV---EVLAKAAKEML 361
Db      840  LKEFTSAVDGAENQARMQETI 862

```

RESULT 11	
_HELPY	
CGA2_HELPY	STANDARD;
	PRT; 1182 AA

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN).
GN CACA OR CAL
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group.
OC Helicobacter.
OX NCBI_TaxID:210;

RP SEQUENCE FROM N.A. / 84-183;
RC STRAIN-ATCC 53726;
RX MEDLINE-9323281; PubMed-8478065;
RA Tsumura M.K.R., Cover T.L., Blaaser M.J.,
RT "Cloning and expression of a high-molecular-mass major antigen of
RT Helicobacter pylori: evidence of linkage to cytotoxin production.",
RL Infect. Immun. 61:1799-1809(1993).
CC -1- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT
OR FUNCTION OF THE CYTOTOXIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib.ch).

DR EMBL, L11714.1; NOT_ANNOTATED_CDS

	878	885	POLY-ASN.	
KM	Antigen.			
FT	DOMAIN			
SO	SEQUENCE	1182 AA;	131503 MM;	C916817E2EE57BB4 CRC64;

Query Match	8.8%	Score 158;	DB 1;	Length 1182;
Best Local Similarity	23.8%	Pred. NO. 1.9;		
Matches 101; Conservative	61;	Mismatches 140;	Indels 122;	Gaps 20

```

0Y 24 PNLFEIS-----KKTESNAVLKVEYETLSDIDELATYATKIGGNLLEA 72
    ||::-----||::||::||::||::||::||::||::||::||::||
Db 541 PNLNLNLTATVVRDRDLEKLIAGLSPOEANKLYKDLSSNKELVGCAL-----NFKAV 553S
OY 73 NQSKNTSLSGAY-----AISDLIAEKLNLVNLKEBELKIDTAKOCSTEFNTKLSEHAV 1277
    ::||::||::||::||::||::||::||::||::||::||::||::||
Db 596 AEAKYT-----GNYDEVRKQKDL--EK--SLKREHLEKGVAAKNLESKSGNKKMKAEK- 646E
OY 128 LGJLNLITDDNQR-----AIIKHKANKD-----KG-----AALEKLFRAVEYLSK 168E
    ||::||::||::||::||::||::||::||::||::||::||::||

```

```

Db      647  -----AQANSQKXEIIFALINKEANRPARALAYQNLKGIKRELSDKLEININDGLDFSK 700
Qy      169  A-----ADDT--LKNAKEL-----TSPTVHNNNSKGD 195
Db      701  SFDGFKNGKNKDFSKAEETLKLGVSKVGDGINDEWISKEVNLMAALNEKKNRKNDFSK 760
Qy      196  NTSANSADK-SVKGPNTLE-ISKKITESNAVLAVK-----ELETLLASIDELATRAIG 247
Db      761  VTKAKSDDENSIKDVIINOKITDKVDELNOAVSVAKIACDFSGVEQALADLKNFSKEOLA 820
Qy      248  KKIOONGL-----AVEAGHNGTLLGAVYIISLITOKLDGJLKNSEKLEKTEENAK 298
Db      821  QOAOQNSFENNGKSEIYIOSVANGVNGTLVGNG-----LSGIATLAKNFSDIKK 870
Qy      299  KCSDEFTFKLEGEHAOLGIEN-----VTDENAKKAILITDAKDCKGALELEKLFKAVENTL 353
Db      871  ELNEKF--KNFNNNNNNNLKNNGEPIYAOVKKTKGVASPEEPIYAOVAKKVTYKKIDOL 928
Qy      354  AKAA 357
Db      929  NOAA 932

```

RESULT	12
M6_STRPY	.
ID M6_STRPY	STANDARD;
780.000	PRT; 483 AA

AC 000003,
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE M PROTEIN, SEROTYPE 6 PRECURSOR.
CN

OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86111835; PubMed=3511046;
RA Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT "Complete nucleotide sequence of type 6 M protein of the group A
Streptococcus. Repetitive structure and membrane anchor.";
RL J. Biol. Chem. 261:1677-1686(1986).

KM SEQUENCE OF 43-122 FROM N.A.
 RP MEDLINE=85166224; PubMed=30883219;
 RX SCOTT J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
 RA "Relationship of M protein genes in group A streptococci."
 RT Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
 CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
 CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
 CC PHAGOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
 CC -1- SIMILARITY: TO OTHER M PROTEINS.
 CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
 CC IN THE REGION OF THE MEMBRANE ANCHOR.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL; M1138; AAA26920.1; -.
DR PIR; A26297; A26297.
DR InterPro; IPR001899; Gram_pos_anchor
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF007376; Gram_pos_anchor; 1.
DR Pfam; PF02340; M_9.
DR PRINTS; PR00015; GPDSANCHOR.

DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
 KW Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
 FT Transmembrane; Coiled coil; Signal.
 FT SIGNAL 1 42
 FT CHAIN 43 483 M PROTEIN, SEROTYPE 6.
 FT DOMAIN 43 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 478 457 MEMBRANE ANCHOR.
 FT DOMAIN 478 483 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 69 138 10 x 7 AA TANDDEM REPEATS.
 FT DOMAIN 157 269 4.5 x 25 AA TANDDEM REPEATS.
 FT DOMAIN 279 347 TWO DIRECTLY REPEATED 27 AMINO ACID
 FT BLOCKS SEPARATED BY 15 AMINO ACIDS.
 FT HYPOTHETIC.
 FT DOMAIN 348 411 GLY/PRO-RICH (CELL WALL-SPANNING).
 FT DOMAIN 442 448 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
 FT PROTEINS.
 FT SEQUENCE 483 AA; 53472 MW; 68f87f28db534448 CRC64;
 SO
 Query Match 8.8%; Score 157.5; DB 1; Length 483;
 Best Local Similarity 24.8%; Pred. No. 0.75;
 Matches 109; Conservative 68; Mismatches 150; Indels 113; Gaps 24;
 OY 1 MACNNSG-----KDGASANSADSVKGP-----NLTEISKTTESNAV-----V 40
 DB 1 MAKNTNRRIVSLRKILKGTASAVA-LSYIGGLVNTNEVSARFPGTVEPNPKAREL 59
 OY 41 LAKEVET--LLASIDELATKAIGKIGNGGLEANSQNTSLSGAVASIDLIAEKLAVL 98
 DB 60 LKRYDVENSMLQANDKLTLE-----NNNL-TDQNNK-----LTENKMLT 99
 OY 99 ---KN--EELKEKIDTAQCSTEFNTKLSEHAVIAGLDNDDDNQAAILKKHANKDKGA 153
 DB 100 DONKMLTEENKMLTDQNNKLTLE--NKLAEKAE---NRLTTEN--KGLTKLSEAESEA 151
 OY 154 AELEKLFK-AVENLSKAADITLKNV--KELTS-----PIVHGNSGKD 194
 DB 152 ANKEHENNEALIGTLKLTIDEYVKDKIAKEQESKETIGTLKLTIDEYVDKIAKEQESKET 211
 OY 195 GNTSANSADSVKGPNLTEISKTTESNAVAVLAKET--ETLLASI----- 238
 DB 212 IGTLLKLTIDEYVK-----DKIAKE--QESKETIGTLKLTIDEYVKDKIAKEQSKODIGALK 266
 OY 239 DELATKAIGKIQONGGLAVEAGHNGTLLAGAYTISKI--ITOKIDGLKNSKK----- 289
 DB 267 QELAKKDGCKNVSEASRKGRLRRDDASREAKQVEKEDLANITAEIDYKKEKQISDSARQ 326
 OY 290 -LKEKTEENAKKCSDEFTKLEGEHAOLGIEVNTDENAKKAILITDAKDKGAEELEKLFK 348
 DB 327 GLRRDLASREAKQVEKALEEANSKLAALKEKLEESKKLTLEKKEK---AELOAKLE 382
 OY 349 AVENLAKAKEMLANSVKEL 368
 DB 383 AE---AKALKLEOLAKQAEEL 399
 RESULT 13
 RBP2_PLAVB
 ID RBP2_PLAVB STANDARD; PRT; 1251 AA.
 AC Q00799;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
 GN RBP2.
 OS Plasmodium vivax (strain Belém).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=31273;
 RN [1]
 RP MEDLINE=92315338; PubMed=1617731;
 RA Gallinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
 RT "A reticulocyte-binding protein complex of Plasmodium vivax

FT mercozoites.";
 RL Cell 69:1213-1226(1992).
 CC -I- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
 CC HUMAN RETICULOCYTE CELLS.
 CC -I- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC (C)
 DR EMBL: M88098; AAA29744.1; -
 KW Malaria; Receptor; Membrane.
 FT NON_TER 1 1
 FT NON_TER 1251 1251
 FT SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;
 SO
 Query Match 8.8%; Score 157.5; DB 1; Length 1251;
 Best Local Similarity 21.3%; Pred. No. 2.1;
 Matches 102; Conservative 72; Mismatches 147; Indels 157; Gaps 22;
 OY 9 DGNASANSADSVK-----GPNLTEISKTTESNAVAVLAV-----KEVETLLASIDELAT 58
 DB 671 EGHGNVAGQVLENIMEIWDENNLSILKQATGKNEIEQIKIHSTLKNNAKIIILGHVDSIA- 729
 OY 59 KAIGKKT-----GNGLSEA-NQSKNTSLSGAVASIDLI 91
 DB 730 KYVGKIKITPELATELLGDALKTAQELKFEKSKNNVLELENNKNTNEIDVHKNIQDAY 789
 OY 92 AEKLNVLKNEELKEKIDTAQOCST-----EPIFNKTKS----- 123
 DB 790 KVALEILLAH---SPEIDTKQKSSKLLKEMGNQIYLKAVVLINQTKNKISSIKSEEAHSV 846
 OY 124 -----EHAVLG-----IDNLTDDNQAAILKKHANKDKGAEELE 157
 DB 847 IGVNSKHSLSLITCSKSDSYDNITALEKQTELQNLANSFTQE---KTNNSD---SKLE 900
 OY 158 KLFPAVENLSKAADITLKNVKELTSPVHGNSGKNTSANSADSVKGPMLTEISK 217
 DB 901 KI-----KTFESLKNLKTLEGE--VVALKASSDNHSHVQSKSPV--NPALSEIEKE 950
 OY 218 ITESNAVAVLAKET--ETLLASIDELATKAIGKIQONGGLAVEAGHNGTLLAGAYTI 273
 DB 951 ETIDISLNTALDELKKGKRCVEVSRYKLIQDYTKELISDTLE-----INTI 997
 OY 274 SKLITQKLDGLKNSKIKKEKTEENAKKCSDEFTKLEGEHAOLGIE--NVTDENAKKAAILI 331
 DB 998 EKVKAKVLAAYIK--KNEDTVQDVLTLNHFNTKOVSNHPTNFDSKNSSEELTKA--V 1053
 OY 332 TDA-----AKDKG-----AAELEKLFKAVENLAKAKAKEN--LANSVK 366
 DB 1054 TDSKTIITSLKGLVIEEVNENTENMTLESSAKTEALYINELKKNKTLNLEYQTSNEVK 1111
 RESULT 14
 MST2_DROHY
 ID MST2_DROHY STANDARD; PRT; 1391 AA.
 AC Q08656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE AXONEME-ASSOCIATED PROTEIN MST101(2).
 GN MST101(2).
 OS Drosophila hydei (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7224;
 RN [1]

Db 1108 VADIKEDLNERRVYDEMOMQOLAKREBELTQTLIRIDESATKATAQAKQBELESQIAEI 1167
QY 67 NNGLEAQSNTSLISGAYASIDLIAEKLVNKEELKEKIDTAKOCSTEFNKIASEHA 126
1168 QEDLEAEKAAARAK----AEKVRDLSEELKALKNELIDSLDTAAQ--OELRSKREQELA 1221
QY 127 VLGIDNLTDDNAQRAIILKKHAN-----KDKGAELEKLFKAVENLSKA-----AOD 172
Db 1222 TL-----KKSLEETVNHGVLADMRRKHSOELINSINDOLENLRKAKTYLEKAKG 1271
QY 173 TLKNAVKELTSPYHGNNSGKDGNTSANSADSVKGPMLTEISKITESNAVLAKEIE 232
Db 1272 TLEENADLATELSVNSSROENDRRKQAESQI-----AEOYKIAEIERARSELQEK 1326
QY 233 TLASIDELATKAIGKIQONGGLAVEAGHN--GTLLAGAYTISKLTOKIDGLN---- 286
Db 1327 TKIQEAEINITNQL-EBAELKASAAVKSASNMESQITEAQOLLEETROKL-GLSSKLRQ 1384
287 ----SEKLEKIEENAKKSEDEFTKKLEGEHNOIGIENTDENAKKAILITDAKDKGAEE 342
Db 1385 IESEKALQEOLEEDDEKRRYERKLAEVTTOM-----OEIKKAEEDADLAK-----E 1433
QY 343 LE---KLFKAVENLAKAKEMLANIS 364
Db 1434 LEEGKKRINKDIEALERQVKELIAQN 1459

Search completed: March 18, 2002, 10:11:57
Job time: 975 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:55 ; Search time 124.19 Seconds
(Without alignments)
433.434 Million cell updates/sec

Title: US-09-596-746A-42
 Perfect score: 1798
 Sequence: 1 MACNNSGKDGNASANSADES.....AVENLAKAKEMLANSVKEL 368

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
 Total number of hits satisfying chosen parameters: 473505

```
Maximum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	901	50.1	192	2	Q9R852	Q9r512 borrellia bu
2	901	50.1	209	2	Q44717	Q44717 borrellia bu
3	899	50.0	191	2	Q9S3P1	Q9s3p1 borrellia bu
4	889	49.4	192	2	Q31117	Q31117 borrellia bu
5	876	48.7	193	2	P94234	P94234 borrellia bu
6	876	48.7	194	2	Q31122	Q31122 borrellia bu
7	876	48.7	201	2	Q9R0R8	Q9r0r8 borrellia bu
8	863	48.0	191	2	Q44726	Q44726 borrellia bu
9	841	46.8	185	2	Q31123	Q31123 borrellia bu
10	826	45.9	178	2	Q44895	Q44895 borrellia bu
11	823	45.8	178	2	Q44879	Q44879 borrellia bu
12	799	41.1	178	2	Q44594	Q44594 borrellia bu
13	653.5	36.3	212	2	Q44705	Q44705 borrellia bu
14	647.5	36.0	193	2	Q31115	Q31115 borrellia bu
15	646	35.9	202	2	Q08831	Q08831 borrellia bu
16	646	35.9	211	2	Q44720	Q44720 borrellia bu
17	644.5	35.8	192	2	Q9S3P3	Q9s3p3 borrellia bu
18	644.5	35.8	192	2	Q9R7B1	Q9r7b1 borrellia bu
19	644.5	35.8	193	2	Q9R853	Q9r853 borrellia bu

20	644.5	35.8	200	2	09R0R9	Q9TqG9	borrella	bu
21	642	35.7	193	2	P94Z37	P96572	borrella	bu
22	641	35.7	188	2	09XDM3	09Xdm3	borrella	bu
23	640.5	35.6	191	2	031120	031120	borrella	bu
24	640	35.6	184	2	09S504	09S504	borrella	bu
25	640	35.6	193	2	P94Z42	P94Z42	borrella	bu
26	640	35.6	211	2	044977	044977	borrella	bu
27	639.5	35.6	192	2	09S3P2	09S3P2	borrella	bu
28	639.5	35.6	210	2	044719	044719	borrella	bu
29	637.5	35.5	191	2	P70818	P70818	borrella	bu
30	637.5	35.5	201	2	P65782	P65782	borrella	bu
31	635	35.3	203	2	008138	008138	borrella	bu
32	634.5	35.3	201	2	P66573	P66573	borrella	bu
33	634	35.3	181	2	034124	034124	borrella	bu
34	630	35.0	194	2	031114	031114	borrella	bu
35	629.5	35.0	201	2	P96514	P96514	borrella	ja
36	629.5	35.0	201	2	P96516	P96516	borrella	ja
37	629	35.0	200	2	09XDM4	09Xdm4	borrella	bu
38	628.5	35.0	182	2	09R7B2	Q9T7B2	borrella	bu
39	627	34.9	211	2	044976	044976	borrella	bu
40	626	34.8	184	2	031121	031121	borrella	bu
41	625	34.8	193	2	P94Z36	P94Z36	borrella	bu
42	624.5	34.7	203	2	050619	050619	borrella	bu
43	623.5	34.7	203	2	050624	050624	borrella	af
44	623.5	34.7	212	2	09K1M5	09K1m5	borrella	af
45	618.5	34.4	201	2	P96571	P96571	borrella	ja

ALIGNMENTS

Query Match	50.1%;	Score 901;	DB 2;	Length 192;
Best Local Similarity	98.9%;	Pred. No. 2.8e-35;		
Matches 184;	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0
QY 1	MACNNSGKDGNSASASAESEYKGPMLTETSKITTESNAVAVLVAVEYETLLASIDELATKA	60		
Db 7	ISCSNNSGKDGNSASASAESEYKGPMLTETSKITTESNAVAVLVAVEYETLLASIDELATKA	66		
OY 61	IGKRTGGNNGLEPNOSKRNSTLSGATAAISDLIAEKLVNLKNEELKEKIDPAKOCSTFTTK	120		
Db 67	IGKRTGGNNGLEPNOSKRNSTLSGATAAISDLIAEKLVNLKNEELKEKIDPAKOCSTFTTK	126		
OY 121	LKSEHVAIVGLDILDTQDNQORALIKKHANKDKGAEELEKLFKAVENLSKAADTLKNAVKE	180		

DB 127 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 186
 QY 181 LTSPIV 186
 DB 187 LTSPIV 192

RESULT 2

044717 PRELIMINARY; PRT; 209 AA.

ID 044717
 AC 044717
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE OSPC.
 GN OSPC.
 OR Borrelia burgdorferi (Lyme disease spirochete).
 NCBI_TaxID=139;
 NCBI_TaxID=139;
 RN 1
 RP SEQUENCE FROM N.A.
 RC STRAIN=NA0;
 RX MEDLINE=94314484; Pubmed=8039931;
 RA Stevenson B., Bockenstedt L.K., Barthold S.W.;
 RT "Expression and gene sequence of outer surface protein C of Borrelia
 burgdorferi reisolated from chronically infected mice."
 RL Infect. Immun. 62:3568-3571(1994).
 DR EMBL: U04240; AAC45538.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 209 AA; 22393 MW; 3707A47DA736FCA CRC64;

Query Match 50.1%; Score 901; DB 2; Length 209;
 Best Local Similarity 98.9%; Pred. No. 3.1e-35;
 Matches 184; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNSGKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 60
 DB 17 ISCNNSKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 76
 QY 61 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120
 DB 77 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 136
 DB 121 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 180
 DB 137 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 196
 QY 181 LTSPIV 186
 DB 197 LTSPIV 202

RESULT 3

0953P1 PRELIMINARY; PRT; 191 AA.

ID 0953P1
 AC 0953P1
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OR Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=28691;
 RX MEDLINE=96296448; Pubmed=8709845;
 RA May I., Gibbs C.P., Schuster R., Dornier F.;
 RT "Evidence for lateral transfer and recombination in OspC variation in

RT Lyme disease Borrelia."
 RL Mol. Microbiol. 18:257-269(1995).
 DR EMBL: I42894; AAB37002.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 191 AA; 20340 MW; 8CA1A64CF17AED8F CRC64;

Query Match 50.0%; Score 899; DB 2; Length 191;
 Best Local Similarity 100.0%; Pred. No. 3.5e-35;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CANSKGDGNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKAIG 62
 DB 1 CANSKGDGNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKAIG 60
 QY 63 KKGNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 122
 DB 61 KKGNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120
 QY 123 SEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 182
 DB 121 SEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 180
 QY 183 SPV 186
 DB 181 SPV 184

RESULT 4

031117 PRELIMINARY; PRT; 192 AA.

ID 031117
 AC 031117
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSPC.
 OR Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=OC7;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF029866; AAB86549.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 192 AA; 20684 MW; 1F0C6BB66E291F6B CRC64;

Query Match 49.4%; Score 889; DB 2; Length 192;
 Best Local Similarity 97.8%; Pred. No. 1e-34;
 Matches 182; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACNSGKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 60
 DB 7 ISCNNSKDNASANSADSVKGNLTETSKITTESNAVYLAKEVEETLLASIDELATKA 66
 QY 61 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 120
 DB 67 IGGKIGNNGLEANOSKNTSLSGAYAISDLIAEKLNVKNEELKEKIDTAQKOSTEFTNK 126
 QY 121 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 180
 DB 127 LKSEHAVLGIDNLTDDNAQRAILKKHANKDKGAELKLFRAVENLSKAAODTLKNAVKE 186

OY 181 LTSPV 186
 DB 187 LTSPV 192

RESULT 5

P94234 PRELIMINARY: PRT: 193 AA.

AC P94234;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.

RT MEDLINE=96296448; PubMed=8709845;
 RT Lively I., Gibbs C.P., Schuster R.,
 RT "Evidence for lateral transfer and recombination in ospC variation in
 RT Lyme disease Borrelia."
 RL MOL. Microbiol. 18:257-269(1995).
 DR EMBL; L42895; AAB37003.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT SEQUENCE 193 193
 SQ SEQUENCE 193 AA: 05868720F061E2A0 CRC64;

Query Match 48.7%; Score 876; DB 2; Length 193;
 Best Local Similarity 100.0%; Pred. No. 4.1e-34;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 NSNGKDGNTSANSADSVKPNLTETISKRTTESNAVLAVKEIETLLASIDELATKAIGK 248
 DB 2 NSNGKDGNTSANSADSVKPNLTETISKRTTESNAVLAVKEIETLLASIDELATKAIGK 61
 OY 249 KIQONGGLAVEAGHNGTLLAGAYTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 308
 DB 62 KIQONGGLAVEAGHNGTLLAGAYTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 121
 OY 309 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 368
 DB 122 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 181

RESULT 6

O31122 PRELIMINARY: PRT: 194 AA.

AC O31122;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=OC12;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RA Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF029871; AAB6554.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.

FT NON_TER 1 1
 FT NON_TER 194 194
 SQ SEQUENCE 194 AA: 20640 MW: 00A5E6E2D2CE0F7E CRC64;

Query Match 48.7%; Score 876; DB 2; Length 194;
 Best Local Similarity 100.0%; Pred. No. 4.1e-34;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 NSNGKDGNTSANSADSVKPNLTETISKRTTESNAVLAVKEIETLLASIDELATKAIGK 248
 DB 10 NSNGKDGNTSANSADSVKPNLTETISKRTTESNAVLAVKEIETLLASIDELATKAIGK 69
 OY 249 KIQONGGLAVEAGHNGTLLAGAYTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 308
 DB 70 KIQONGGLAVEAGHNGTLLAGAYTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 129
 OY 309 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 368
 DB 130 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 189

RESULT 7

O9ROR8 PRELIMINARY: PRT: 201 AA.

AC O9ROR8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid cp26.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=3B6;
 RX MEDLINE=20002545; PubMed=10531219;
 RA Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
 RT "Population dynamics of a naturally occurring heterogeneous mixture of
 RT Borrelia burgdorferi clones."
 RL Infect. Immun. 67:5709-5716(1999).
 DR EMBL; AF074465; AAD23912.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 KW plasmid.
 FT NON_TER 201 201
 SQ SEQUENCE 201 AA: 21457 MW: 806F198295101B07 CRC64;

Query Match 48.7%; Score 876; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 4.3e-34;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 NSNGKDGNTSANSADSVKPNLTETISKRTTESNAVLAVKEIETLLASIDELATKAIGK 248
 DB 20 NSNGKDGNTSANSADSVKPNLTETISKRTTESNAVLAVKEIETLLASIDELATKAIGK 79
 OY 249 KIQONGGLAVEAGHNGTLLAGAYTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 308
 DB 80 KIQONGGLAVEAGHNGTLLAGAYTTISKLTQKLDGLKNSKLEKEIKENAKKCSDFPKKL 139
 OY 309 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 368
 DB 140 EGEHAOLGIENVTDENAKKAILITDAKDKGALELEKFAVENLAKAKEMLANSVKEL 199

RESULT 8

O44726 PRELIMINARY: PRT: 191 AA.

AC O44726;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
 GN OSCP.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid cp26.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RX MEDLINE=94314437; PubMed=8039891;
 RA Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;
 RT "Humoral immune response to outer surface protein C of Borrelia
 burgdorferi in Lyme disease: role of the immunoglobulin M response in
 the serodiagnosis of early infection."
 RL Infect. Immun. 62:3213-3221(1994).
 DR EMBL: U08284; AAA21460.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 Pfam: PF01441; Lipoprotein_6; 1.
 Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 SQ SEQUENCE 191 AA; 20150 MW; C49AA030F0A28717 CRC64;

Query Match 48.0%; Score 863; DB 2; Length 191;
 Best Local Similarity 99.4%; Pred. No. 1.6e-33;
 Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 190 NSGKGNTSANSADSVGPNLTETSKKITESNAVLAKEITLLASIDELATKAIGK 249
 DB 1 NSGKGNTSANSADSVGPNLTETSKKITESNAVLAKEITLLASIDELATKAIGK 60
 QY 250 IQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLE 309
 DB 61 IQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLE 120
 QY 310 GEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKAKEMLANSVKEL 368
 DB 121 GEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKAKEMLANSVKEL 179

RESULT 9
 ID 031123 PRELIMINARY; PRT; 185 AA.
 AC 031123;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC NCBI_TaxID=139;
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OC13;
 RA Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF029872; AAB86555.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 Pfam: PF01441; Lipoprotein_6; 1.
 Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 SQ SEQUENCE 185 AA; 19673 MW; 58D6FE3C7769CAF CRC64;

Query Match 46.8%; Score 841; DB 2; Length 185;
 Best Local Similarity 97.7%; Pred. No. 1.6e-32;
 Matches 172; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 189 NSGKGNTSANSADSVGPNLTETSKKITESNAVLAKEITLLASIDELATKAIGK 248

DB 10 NSGKGNTSANSADSVGPNLTETSKKITESNAVLAKEITLLASIDELATKAIGK 69
 QY 249 KIQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKL 308
 DB 70 KIQQNGGLAVEAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKL 129
 QY 309 EGEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKAKEMLANS 364
 DB 130 EGEHQAOLGIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKAKEMLANS 185

RESULT 10
 ID 044995 PRELIMINARY; PRT; 178 AA.
 AC 044995;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MUL;
 RX MEDLINE=95286481; PubMed=7768799;
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 Hansen K.;
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC."
 RL J. Bacteriol. 177:3036-3044(1995).
 DR EMBL: X84779; CA59250.1;
 DR InterPro: IPR001800; Lipoprotein_6.
 Pfam: PF01441; Lipoprotein_6; 1.
 Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 178 AA; 18911 MW; EE50CE48EADF1CA CRC64;

Query Match 45.9%; Score 826; DB 2; Length 178;
 Best Local Similarity 98.8%; Pred. No. 7.8e-32;
 Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 196 NTSANSADSVGPNLTETSKKITESNAVLAKEITLLASIDELATKAIGKIOQNG 255
 DB 1 NTSANSADSVGPNLTETSKKITESNAVLAKEITLLASIDELATKAIGKIOQNG 60
 QY 256 LAYAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLEGEHAOL 315
 DB 61 LAYAGHNGTLLAGAYTISKLTQKLDGKNSKLEKIEKAKKSEDFTKLEGEHAOL 120
 QY 316 GIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKAKEMLANSVKEL 368
 DB 121 GIEVNTDENAKKAILITDAKDKGAELKLFRAVENLAKAKEMLANSVKEL 173

RESULT 11
 ID 044979 PRELIMINARY; PRT; 178 AA.
 AC 044979;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=272;
 RX MEDLINE=95286481; PubMed=7768799;

RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC."
 RL J. Bacteriol. 177:3036-3044(1995).
 DR EMBL: X84783; CAAS9253.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 178 178
 SQ SEQUENCE 178 AA; 18938 MW; 972974753EFEB1E9 CRC64;

Query Match 45.8%; Score 823; DB 2; Length 178;
 Best Local Similarity 98.3%; Pred. No. 1.1e-31;
 Matches 170; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 196 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 255
 DB 1 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 60
 DB 256 LAVEAGNGTLLAGAYTTSKLTQKLDGLKNSKEKIEENAKKCSDEFTKKLEGEHAOL 315
 DB 61 LAVEAGNGTLLAGAYTTSKLTQKLDGLKNSKEKIEENAKKCSDEFTKKLEGEHAOL 120
 OY 316 GIEVNTDENAKKAILITDAARDKGALEKLEKFAVENLAKAKEMLANSVKEL 368
 DB 121 GIEVNTDENAKKAILITDAARDKGALEKLEKFAVENLAKAKEMLANSVKEL 173

RESULT 12
 O44994 PRELIMINARY; PRT: 178 AA.

AC O44994;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS OSPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1P;
 RX MEDLINE=95286481; PubMed=7768799;
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
 RA Hansen K.;
 RT "Evolution of the Borrelia burgdorferi outer surface protein OspC."
 RL J. Bacteriol. 177:3036-3044(1995).
 DR EMBL: X84782; CAAS9253.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 178 178
 SQ SEQUENCE 178 AA; 18818 MW; 79B435D0837DA36D CRC64;

Query Match 41.1%; Score 739; DB 2; Length 178;
 Best Local Similarity 89.0%; Pred. No. 8.4e-28;
 Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 196 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 255
 DB 1 NTSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRAIKKIQONG 60
 OY 256 LAVEAGNGTLLAGAYTTSKLTQKLDGLKNSKEKIEENAKKCSDEFTKKLEGEHAOL 315
 DB 61 LAVEAGNGTLLAGAYTTSKLTQKLDGLKNSKEKIEENAKKCSDEFTKKLEGEHAOL 120
 OY 316 GIEVNTDENAKKAILITDAARDKGALEKLEKFAVENLAKAKEMLANSVKEL 368
 DB 121 GIEVNTDENAKKAILITDAARDKGALEKLEKFAVENLAKAKEMLANSVKEL 173

DB 121 GIEVNTDENAKKAILITDAARDKGALEKLEKFAVENLPOAKEMLANSVKEL 173

RESULT 13
 O44705 PRELIMINARY; PRT: 212 AA.
 AC O44705;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C.
 OS OSPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2591;
 RX MEDLINE=94041630; PubMed=8225587;
 RA Padula S.J., Samperi A., Dias F., Szczepanski A., Ryan R.W.;
 RT "Molecular characterization and expression of p23 (OspC) from a North
 American strain of Borrelia burgdorferi."
 RL Infect. Immun. 61:5097-5105(1993).
 DR EMBL: U01892; AAA16057.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 212 AA; 22270 MW; FB2EF2673A384276 CRC64;

Query Match 36.3%; Score 653.5; DB 2; Length 212;
 Best Local Similarity 72.4%; Pred. No. 9.5e-24;
 Matches 139; Conservative 21; Mismatches 29; Indels 3; Gaps 2;

OY 1 MACNSGKDGKNSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRA 60
 DB 17 ICNNSGKDGKNSANSADSVKGPMLTEISKRTESNAVVLAVKEIETLLASIDELATRA 76
 OY 61 ICKKIGNNGLEANOSKNTSLSGAVAIISDLIAEKLNVLN-BELKKKIDTAQOCSTEFN 119
 DB 77 IGNLIQNLGNAGANONGSLAGAVIISTLIAEKLNLGKLNSEBELKIEDAKCKNAQFD 136
 OY 120 KKKSEHAVIGLDN--LTDNAORATILKHKANKDKGALEKLEKFAVENLAKAKEMLAN 177
 DB 137 KKKSSHAELGIANGAASDANAKAAILIKGTGDKGAOLEKLEKFSYKNSKAAQETLNS 196
 OY 178 VKELTSPYVGN 189
 DB 197 VKELTSPYVAEN 208

RESULT 14
 O31115 PRELIMINARY; PRT: 193 AA.
 AC O31115;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 OS OSPC.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OC3;
 RX Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
 RT Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF029862; AAB86545.1; -;
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR Prodom: PD001149; Lipoprotein_6; 1.
 FT NON_TER 1
 FT NON_TER 193 193

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:34 ; Search time 118.14 Seconds
(without alignments)
351.118 Million cell updates/sec

Title: US-09-596-746A-52

Perfect score: 2750
Sequence: 1 MACNSGKDCQNTSANSNADES.....KAAQALFNSVKELTPVVA 560

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 522463 seqs, 74073290 residues

Number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2750	100.0	560	22	AAB62724
2	1758	63.9	410	22	AAB62740
3	1754	63.8	386	22	AAB62727
4	1729.5	62.9	384	22	AAB62726
5	1726.5	62.8	408	22	AAB62737
6	1493.5	54.2	373	22	AAB62711
7	1488.5	54.1	397	22	AAB62729
8	1487	54.1	374	22	AAB62710
9	1484	54.0	398	22	AAB62728
10	1474	53.6	400	22	AAB62739
11	1469.5	53.4	401	22	AAB62738

Result No.	Score	Query Match	Length	ID	Description
12	1465	53.3	377	22	AAB62713
13	1460.5	53.1	378	22	AAB62712
14	1452	52.8	378	22	AAB62725
15	1444	52.5	401	22	AAB62733
16	1198.5	43.6	369	22	AAB62714
17	1198.5	43.6	369	22	AAB62716
18	1195.5	43.5	393	22	AAB62731
19	1195.5	43.5	393	22	AAB62732
20	1194	43.4	370	22	AAB62715
21	1191	43.3	394	22	AAB62730
22	1167.5	42.5	368	22	AAB62717
23	1162.5	42.3	367	22	AAB62721
24	1159.5	42.2	391	22	AAB62735
25	1159.5	42.2	391	22	AAB62736
26	1158	42.1	368	22	AAB62718
27	1155	42.0	392	22	AAB62734
28	986.5	35.9	466	16	AAR75740
29	928	33.7	192	22	AAB62721
30	928	33.7	210	16	AAW11935
31	928	33.7	210	16	AAR75727
32	926	33.7	192	15	AAR60886
33	922	33.5	587	16	AAR75746
34	920	33.5	209	22	AAB62720
35	915	33.3	466	16	AAR75739
36	908	33.0	192	22	AAB62703
37	889	32.3	211	22	AAB62722
38	889	32.3	212	16	AAR75729
39	886	32.2	212	18	AAW41824
40	885	32.2	206	22	AAB62723
41	882	32.1	207	16	AAR75730
42	880	32.0	194	15	AAR60896
43	878	31.9	207	18	AAW41823
44	864	31.4	176	15	AAR62772
45	858	31.2	212	12	AAR13140

ALIGNMENTS

RESULT 1
ID AAB62724 standard; Protein: 560 AA.
XX
AC AAB62724;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 52.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000MO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNV) UNIV NEW YORK STATE RES FOUND.
XX
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Datwyler RJ, Seino G, Dykhizen D, Luft BJ, Gomes-Solecki M;
DR WPI: 2001-050113/06.
XX N-PSDB: AAR29028.
XX
PT Compositions of ospC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -

XX Claim 43: Page 112-113; 160pp; English.
PS
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 2750; DB 22; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNSGCDGTSANSADESVKGPNTLTSKRTDSNANVLAVKEVELLSIDEIATAKA 60
1 macnsgcdgtsansadesvkgpnltskrtidsnavlavkevelllsideiaaka 60
IGKKIHONGIDTEYNHNGSLAGAYAISTLIKOKLIDGKNEGLKEKIDAAKCSSEPTN 120
61 igkkihongidteynhngslagayaistlikqkldgkneglkeidaakkcseltfn 120
61 igkkihongidteynhngslagayaistlikqkldgkneglkeidaakkcseltfn 120
QY 121 KLAERHTDLGEGYTDADAKPAILKTNGTKTGABELGKLFESYEVLSKAAKEMLANSVK 180
121 klaerhtdlggytydadakpaillktngtktgabelgklfesyevlskaakemlansvk 180
121 klikehtdlggytydadakpaillktngtktgabelgklfesyevlskaakemlansvk 180
QY 181 ELTSPVAESPAMSGNSGKGDSTNPADSAKGPNTLTSKRTDSNANVLAVKEVEY 240
181 eltspvaespamsgnskgdsastnpadesakgpnltskrtidsnavlavkevet 240
QY 241 LVLSIDELAKRAIGOKIDNNNGSLAALNNQNSLLAGAYAISTLTTEKLSKLNLEELKTE 300
241 lvlsidelakraigokidnnngslaalnnqnsllagayaistltteklsklnleelkte 300
QY 301 IAKKAKCSSEPTNKLKSGHADLGKODATDHAKAAILKTHATTGKAEKFDLFESVEGL 360
301 iakkakcsseptnklksghadlgkodatdhakaailkthattgkakefdlfevegl 360
301 iakkakcsseptnklksghadlgkodatdhakaailkthattgkakefdlfevegl 360
QY 361 LKAQVALTNSVKELTGNHNSGDSASTNPDESAGPNTLVTSKRTDSNANVLAVKEVE 420
361 lkaqvaltnsvkeltnhnsghdsastnpdesakgpnlvtskrtidsnavlavkeve 420
361 lkaqvaltnsvkeltnhnsghdsastnpdesakgpnlvtskrtidsnavlavkeve 420
QY 421 ALLSSIDELSKAIGKIKINDGTLDNEANRNESLITAGAYEISKLITOKLSVLNSEELEKKI 480
421 allssidelksaigkikindgtldneanrneslilagayeisklitqkrlsvlnseelkxi 480
421 allssidelksaigkikindgtldneanrneslilagayeisklitqkrlsvlnseelkxi 480
QY 481 KEAKDCSQKFTTKLKDASHAELGIQSVODDNAKKAAILKTHTGDKGAKLELELFKSLSLS 540
481 keakdcsqkfttklkdsaelgiqsvoddnakkaailkthtgdkgaklelelfksls 540
481 keakdcsqkfttklkdsaelgiqsvoddnakkaailkthtgdkgaklelelfksls 540
QY 541 KAAQALTNVSKELTNPVA 560
541 kaaqaaltnsvkeltnpva 560
541 kaaqaaltnsvkeltnpva 560
DB
RESULT 2
AAB62740 ID AAB62740 standard; Protein; 410 AA.
XX
AC AAB62740;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 84.
XX
KW Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
XX
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.

XX
PD 28-DEC-2000.
XX
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
XX
PR 18-JUN-1999; 99US-0140042.
XX
XX
PA (UNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M,
XX
XX WPI; 2001-050113/06.
DR N-PSDB; AAF29044.
XX
XX
PT Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX
PS Claim 43: Page 155-156; 160pp; English.
CC
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 410 AA;

Query Match 63.9%; Score 1758; DB 22; Length 410;
Best Local Similarity 95.1%; Pred. No. 1.8e-109;
Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

QY 188 AESPAMSGNSGKGDSTNPADSAKGPNTLTSKRTDSNANVLAVKEVELLSIDE 247
188 aespamsgnskgdsastnpadesakgpnltskrtidsnavlavkevelllside 247
20 aespamsgnskgdsastnpadesakgpnltskrtidsnavlavkevelllside 247
QY 248 LAKKAIGOKIDNNNGSLAALNNQNSLLAGAYAISTLTTEKLSKLNLEELKTEIATAKAC 307
248 lakkaigokidnnngslaalnnqnsllagayaistltteklsklnleelkteiatakkc 307
248 lakkaigokidnnngslaalnnqnsllagayaistltteklsklnleelkteiatakkc 307
QY 308 SEEFNTKLKSGHADLGKODATDHAKAAILKTHATTGKAEKFDLFESVEGLKAAQVA 367
308 seefntklksghadlgkodatdhakaailkthattgkakefdlfeveglkkaqva 367
308 seefntklksghadlgkodatdhakaailkthattgkakefdlfeveglkkaqva 367
QY 368 LTNVSKELG-----RNSGDSASTNPDESAGPNTLVTSKRTDSNANVLAVKEVE 415
368 ltnsvkelg-----rnsghdsastnpdesakgpnlvtskrtidsnavlavkeve 415
368 ltnsvkelg-----rnsghdsastnpdesakgpnlvtskrtidsnavlavkeve 415
QY 416 VKEVEALSSIDELSKAIGKIKINDGTLDNEANRNESLITAGAYEISKLITOKLSVLNSEE 475
416 vkevealssidelksaigkikindgtldneanrneslilagayeisklitqkrlsvlnsee 475
416 vkevealssidelksaigkikindgtldneanrneslilagayeisklitqkrlsvlnsee 475
QY 476 LKKRIKEAKDCSQKFTTKLKDASHAELGIQSVODDNAKKAAILKTHTGDKGAKLELELFKS 535
476 lkkrikeakdcsqkfttklkdsaelgiqsvoddnakkaailkthtgdkgaklelelfks 535
476 lkkrikeakdcsqkfttklkdsaelgiqsvoddnakkaailkthtgdkgaklelelfks 535
QY 536 LESISKAQAALTNVSKELTNPVA 560
536 lesiskaqaaltnsvkeltnpva 560
536 lesiskaqaaltnsvkeltnpva 560
DB 380 lesiskaqaaltnsvkeltnpva 404
RESULT 3
AAB62727 ID AAB62727 standard; Protein; 386 AA.
XX
AC AAB62727;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 58.
XX

QY 422 LLSIDELSKAIGKIRKNDGFLDNEANRNESLIGAVEISKLITOKLSVNSEELKKIK 481
|||||
Db 240 llsidelskaigkirkndgfloneanrnesliagayeisklitqklsvlnseelkkik 299
QY 482 EAKDCSOKFTTKLKDShAELIGQSVODDNARKKAILKTHGTRDKGAKLEELFKSLLESLSK 541
|||||
Db 300 eakdcsgkfttklkdshaeligqsvoddnarkkailkthgtrdkgakleelfkleslsk 359
QY 542 AAOAALTNYSVKELTNPYVA 560
|||||
Db 360 aagaaltnsvkeltlnpyva 378
RESULT 5
AAB62737 standard; Protein: 408 AA.
ID AAB62737
AAB62737;
DT 03-APR-2001 (first entry)
XX DE Borrelia sp chimeric ospc protein SEQ ID NO: 78.
XX KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX OS Chimeric - Borrelia sp.
XX OS Chimeric - Borrelia sp.
XX PN WO200078966-A1.
XX PD 28-DEC-2000.
XX PF 19-JUN-2000; 2000WO-US16915.
XX PR 18-JUN-1999; 99US-0140042.
XX PA (UYNV) UNIV NEW YORK STATE RES FOUND.
XX PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29041.
XX PT Compositions of Ospc polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX PS Claim 43; Page 147-148; 160pp; English.
XX CC The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads
XX CC to Lyme disease.
XX SO Sequence 408 AA;
Query Match 62.88; Score 1726.5; DB 22; Length 408;
Best Local Similarity 66.7%; Pred. No. 2.2e-107;
Matches 373; Conservative 2; Mismatches 3; Indels 181; Gaps 1;
QY 2 ACNNSGDKGNTSANSASDESVCNPLTEISKKTDSNVLAVKFEVALLSIDELAAKAI 61
:|||||
Db 25 scnnsydgntsansadesvcpnlteiskktidsnvlavkeevallsidelaakai 84
QY 62 GKRIHQNNGDTEYNHNGSLAGAYASTLIKQKIDGKNGELKEKIDPAKCESETFNK 121
|||||
Db 85 gkrihgngdteynhngslagayastlikqkldgkngelkrekidaakcetfnk 144
QY 122 LKEKHTDLGKEGVTDADAKKAIKTNKTKGAEELGKLESEVLSKAKEMLANSYKE 181
|||||

Db 145 lkehtdlgkegvtdadakeailktngtktkgaeeigklfesvevlskaakemlansyke 204
QY 182 LTFSPVAESPMGNSGCGDSASTNPADESAGKPNLTEISKKTDSNAFVLAKEVETL 241
|||||
Db 205 ltfspvaeasp----- 214
QY 242 VLSDIELAKKAIIGKIDNNGLAALNNONGSLAGAYASTLITEKLSKLKNEELKTEI 301
Db 215 ----- 214
QY 302 AKARKSEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKFKDLFESVEGL 361
Db 215 ----- 214
QY 362 KAAOVALTNYSVKELGHRNNSGCGDSASTNPPDSAGKPNLTIVISKKTDSNAFLAVKEVEA 421
:|||||
Db 215 -----kpfhgnsgdsastnpdsakgpnltiviskktidsnaflavkevea 263
QY 422 LLSIDELSKAIGKIRKNDGFLDNEANRNESLIGAVEISKLITOKLSVNSEELKKIK 481
|||||
Db 264 llsidelskaigkirkndgfloneanrnesliagayeisklitqklsvlnseelkkik 323
QY 482 EAKDCSOKFTTKLKDShAELIGQSVODDNARKKAILKTHGTRDKGAKLEELFKSLLESLSK 541
|||||
Db 324 eakdcsgkfttklkdshaeligqsvoddnarkkailkthgtrdkgakleelfkleslsk 383
QY 542 AAOAALTNYSVKELTNPYVA 560
|||||
Db 384 aagaaltnsvkeltlnpyva 402
RESULT 6
AAB62711 standard; Protein: 373 AA.
ID AAB62711
AAB62711;
DT 03-APR-2001 (first entry)
XX DE Borrelia sp chimeric ospc protein SEQ ID NO: 26.
XX KM Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX OS Chimeric - Borrelia sp.
XX OS Chimeric - Borrelia sp.
XX PN WO200078966-A1.
XX PD 28-DEC-2000.
XX PF 19-JUN-2000; 2000WO-US16915.
XX PR 18-JUN-1999; 99US-0140042.
XX PA (UYNV) UNIV NEW YORK STATE RES FOUND.
XX PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX PI Dattwyler RJ, Seimost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29015.
XX PT Compositions of Ospc polypeptides from strains of Borrelia which cause
XX Lyme disease are used to immunize animals and detect immune responses
XX to Lyme disease -
XX PS Claim 43; Page 81; 160pp; English.
XX CC The present invention provides compositions comprising ospc proteins and
XX CC chimeric ospc proteins from members of the Borrelia genus. These may be
XX CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
XX CC vaccines against Borrelia infection, which is spread by ticks and leads

CC	to Lyme disease.
XX	
SQ	Sequence 373 AA;

Query Match	54.28;	Score 1491.5;	DB 22;	Length 373;
Best Local Similarity	82.68;	Pred. No. 9.3e-92;		
Matches 309;	Conservative 21;	Mismatches 43;	Indels 1;	Gaps 1.

QY	2	ACNNNSKDGNTSANSADSESVYKGPULTEISKRITDTSNVAVLAVEVEALLSIDIEIAAKAI	61
			62
			63
			64
			65
			66
			67
			68
			69
			70
			71
			72
			73
			74
			75
			76
			77
			78
			79
			80
			81
			82
			83
			84
			85
			86
			87
			88
			89
			90
			91
			92
			93
			94
			95
			96
			97
			98
			99
			100
			101
			102
			103
			104
			105
			106
			107
			108
			109
			110
			111
			112
			113
			114
			115
			116
			117
			118
			119
			120
			121
			122
			123
			124
			125
			126
			127
			128
			129
			130
			131
			132
			133
			134
			135
			136
			137
			138
			139
			140
			141
			142
			143
			144
			145
			146
			147
			148
			149
			150
			151
			152
			153
			154
			155
			156
			157
			158
			159
			160
			161
			162
			163
			164
			165
			166
			167
			168
			169
			170
			171
			172
			173
			174
			175
			176
			177
			178
			179
			180
			181
			182
			183
			184
			185
			186
			187
			188
			189
			190
			191
			192
			193
			194
			195
			196
			197
			198
			199
			200
			201
</			

RESULT 7
AAB62720

ID AAB62729 standard; Protein; 397 AA.

AC AAB62729

DT 03-APR-2001 (first entry)

DE Borrella sp chimeric ospc protein SEQ ID NO: 62

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

Chimeric - Borrella sp.
Chimeric - Borrella sp.

PN WO200078966-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UYN) UNIV NEW YORK STATE RES FOUND
PA (BROO-) BROOK BIOTECHNOLOGIES INC.

PI Datwyljer RJ, Selnost G, Dykhulzen D, Luft BJ, Gomes-Solecki M;

DR WPI; 2001-050113/06.

DR N-PSDB; AAF29033.

Compositions of OspC polypeptides from strains of *Borrelia* which cause Lyme disease are used to immunize animals and detect immune responses to Lyme disease -

PS Claim 43; Page 125-126; 160pp; English.

CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.

SQ Sequence 397 AA;

Query Match	54.1%	Score 1488.5	DB 22	Length 397
Best Local Similarity	82.4%	Pred. No. 1.6e-91		
Matches 308	22	Mismatches 43	Indels 1	Gaps 1

QY	2	ACNNSKQGNFSANSABESVYGPMLTEISKRTDTSNANVLAVKEVALLSSIDEIAAKAI	61
Dp	25	scmsykgdgnfsansadesvxygpnltelskktcdtsnavllyavkevellssideiaakai	84
QY	62	GKTIHONNGLDTEYHNHNGSLLAGAVASTLIKOKLDLKNELKEKIDAAKCSFTFNK	121
Dp	85	gkltlhngmldteyhngslsllagayastllkqrlldgllkneglktekidaakkcsftfnk	144
QY	122	LKEKHTDYGKGVTDADAKKAILKTNGTKTKGAEELCKLEESVEVLSAAKEMLANSVKE	181
Dp	145	lkehctdlygkegvcdadakeailktngctkcygaeelgkltsevevlskaakemlansvke	204
QY	182	LTSPPVAESPMPGNSNGSGGSASTNPADSAKCPNLTETSKRTDTSNANVLAVKEVELT	241
Dp	205	ltsppvaespmpvngnsygdgnf-sansadesvxygpnltelskktcdtsnavllyavkeletl	263
QY	242	VLSDTELAKKAIKGRKINNNNNLALNNONGSLLAGAVASTLITETKLSKLNLEELKTEI	301
Dp	264	lasdelatetacklgyklqngsllyaveaghnngllagaytclsklltqkrlldgllkneaklkekl	323
QY	302	AKAKKCSSEFTFNKLSGSHADLKGDATDHDHAKAAILKTHTATTDGAKEFKDLFESEVGLL	361
Dp	324	enakkcsedftfckllgeghaaglyenvcdenakkailltldaackdgaaealeklfavenla	383
QY	362	KAAQVALTNSVKEL 375	
Dp	384	kaakemlansvkel 397	

RESULT 8

ID AAB62710 standard; Protein; 374 AA

AC AAB62710;

DT 03-APR-2001 (first entry)

Borrelia sp chimeric ospC protein SEQ ID NO: 24.

KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS	Chimeric - Borrelia sp.
OS	Chimeric - Borrelia sp.

PN WO200078966-A1

PD 28-DEC-2000

PF 19-JUN-2000; 2000WO-US16915.

PR 18-JUN-1999; 99US-0140042.

PA (UUNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO -) BROOK BIOTECHNOLOGIES INC.

PI	Dattwyler RJ, Sehnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M,
XX	WPI; 2001-050113/06.
DR	

DR N-PSDB: AAF29014.
XX Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43; Page 78-79; 160pp; English.
PS
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 374 AA:
SO
Query Match 54.1%; Score 1487; DB 22; Length 374;
Best Local Similarity 83.4%; Pred. No. 1.9e-91;
Matches 312; Conservative 20; Mismatches 40; Indels 2; Gaps 2;
QY 2 ACNNSGKDNSTANSADSVKGNPLTEISKRTTDSNAVLAVKEVALSSIDEIAAKAI 61
Db 1 acnnsykgdgnsansadesvkgpnlteiskrtidsnavlavkevalssideiaakai 60
QY 62 GKRIHONNGDTEYNNHNSLAGAYAI STLKOKIDGLKNGELKEKIDAAKCEFTFNK 121
Db 61 gkrlhngngldteyngnslagayaistlikgldglneglnkeidaakcsetfknk 120
QY 122 LKEKHTDGLKGGVTDADAKKALIKTNGTKTGAELKLFESVEVLSKAKEMLANSVKE 181
Db 121 lkehtldlgkgyvtdadakealiktngtkgaeeilgklfesvevlskaakemlansvke 180
QY 182 LTSPPVAESPMGNSGSGDSASTNPADESAKGPNLTEISKRTTDSNAVLAVKEVETL 241
Db 181 ltsppvaespmwmsngsdydngnt-sansadesvkgpnlteiskrtidesnavlavkevcl 239
QY 242 VLSIDELAKKAIGOKIDNNNGSLAALNNONGSLAGAYAI STLTEKLSKLNLEELKTEI 301
Db 240 ltsidela-kaigkklkndvsladneadngslisgayslslntkksaidselgaei 298
QY 302 AKAKKCEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDGAKFEKDLFESVEGL 361
Db 299 ekakckseefaklkghehtldlgkgyvtdadnakaliktndktkgadelleklfesvknls 358
QY 362 KAAQVALTNSYKEL 375
Db 359 kaakemltnsvkel 372
RESULT 9
AAB62728
ID AAB62728 standard; Protein: 398 AA.
XX
XX AAB62728;
AC
XX
XX 03-APR-2001 (first entry)
DT
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 60.
KW
XX
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS
XX
XX Chimeric - Borrelia sp.
OS
XX
XX Chimeric - Borrelia sp.
PN
XX
XX WO200078966-A1.
XX
XX 28-DEC-2000.
PD
XX
XX 19-JUN-2000; 2000MO-US16915.
PF
XX
XX 18-JUN-1999; 99US-0140042.
PR
XX

PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX Datwyler RJ, Seinoest G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PI
XX
XX WPI; 2001-050113/06.
DR
XX
XX N-PSDB: AAF29032.
PS
PT Compositions of Ospc polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
XX Claim 43; Page 123; 160pp; English.
PS
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. atzeili or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
XX Sequence 398 AA:
SO
Query Match 54.0%; Score 1484; DB 22; Length 398;
Best Local Similarity 83.2%; Pred. No. 3.2e-91;
Matches 311; Conservative 21; Mismatches 40; Indels 2; Gaps 2;
QY 2 ACNNSGKDNSTANSADSVKGNPLTEISKRTTDSNAVLAVKEVALSSIDEIAAKAI 61
Db 25 scnnsykgdgnsansadesvkgpnlteiskrtidsnavlavkevalssideiaakai 84
QY 62 GKRIHONNGDTEYNNHNSLAGAYAI STLKOKIDGLKNGELKEKIDAAKCEFTFNK 121
Db 85 gkrlhngngldteyngnslagayaistlikgldglneglnkeidaakcsetfknk 144
QY 122 LKEKHTDGLKGGVTDADAKKALIKTNGTKTGAELKLFESVEVLSKAKEMLANSVKE 181
Db 145 lkehtldlgkgyvtdadakealiktngtkgaeeilgklfesvevlskaakemlansvke 204
QY 182 LTSPPVAESPMGNSGSGDSASTNPADESAKGPNLTEISKRTTDSNAVLAVKEVETL 241
Db 205 ltsppvaespmwmsngsdydngnt-sansadesvkgpnlteiskrtidesnavlavkevcl 263
QY 242 VLSIDELAKKAIGOKIDNNNGSLAALNNONGSLAGAYAI STLTEKLSKLNLEELKTEI 301
Db 264 ltsidela-kaigkklkndvsladneadngslisgayslslntkksaidselgaei 322
QY 302 AKAKKCEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDGAKFEKDLFESVEGL 361
Db 323 ekakckseefaklkghehtldlgkgyvtdadnakaliktndktkgadelleklfesvknls 382
QY 362 KAAQVALTNSYKEL 375
Db 383 kaakemltnsvkel 396
RESULT 10
AAB62739
ID AAB62739 standard; Protein: 400 AA.
XX
XX AAB62739;
AC
XX
XX 03-APR-2001 (first entry)
DT
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 82.
KW
XX
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
OS
XX
XX Chimeric - Borrelia sp.
OS
XX
XX Chimeric - Borrelia sp.
PN
XX
XX WO200078966-A1.
XX

Query Match	Best Local Similarity	Matches 307; Conservative 23; Mismatches 43; Indels 4; Gaps 2;
53.6%; Score 1474; DB 22; Length 400;	81.4%; Pred. No. 1.5e-90;	
Sequence 400 AA:		
2 ACNNGSKDGNFSANAGDSVKGPNLTETISKTTDSNAVLAKVEYALSSIDEIAAKAI 61		
25 scnsgkdcgnfsaansadesvkgpnltelskktidsnavllavkeaalssideiaakai 84		
62 GKIKIQQNGLDPEYHNHNSLLAGAAISTLIKQKLDGLKNEELKSKIDAAKCSSETFTYK 121		
85 gkikimgngldpeyhnhngslagayaisltikqkldglneglkexidaakkcssetftnk 144		
122 LKEKHTDLCGEGVTDADAKKALIKTNGTKYKGAELGKLFESVEVLSKAAKEMLANSVKE 181		
145 lkehtldlgkegvtdadakeallkngctkkgaeelgklfesvevlskaakemlansvke 204		
182 LTPSPVAAS---PAMGNSNGSGGSDASINRPADESAAGPMLTETISKTTDSNAFVLAVKEV 238		
205 ltpspvaasppkpsrpnvannsgkdcgn--saansadesvkgpnltelskktidesnavlavkai 263		
239 ETLVLSIDELAKKAIAGKIDNNNGTAAALNNNGSLLAGAAISTLTETLSKTLKLNDEELK 298		
264 etllsidselakkaigkikgngslaveaghnctllagayelskktlctqkldglnsekik 323		
299 TEIAAKKCSSEFTYKLSGSHADLCGODATDDHAKAAILKTHATTDKGAKERKDLFESYE 358		
324 etlienaakkcsedftkkllegehaqglentvdenakkailltcaadkdgaaetleklfave 383		
359 GLKKAQVALTNSVKEL 375		
384 nlakaakemlansvkel 400		
RESULT 11		
AAB62738		
AAB62738 standard; Protein; 401 AA.		
AAB62738;		
03-APR-2001 (first entry)		
Borrelia sp chimeric ospC protein SEQ ID NO: 80.		
Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.		

Query Match	53.4%	Score 1469.5	DB 22	Length 401
Best Local Similarity	82.2%	Pred. No. 3e-90		
Matches 310	Conservative 22	Mismatches 40	Indels 5	Gaps 3
Sequence	401 AA			
Query	2	AANNNGKCGNTANSNSADESVKGNLTETSKRTTDSNAVLAVKEVYALISIDEITAAKAI	61	
Db	25	scnmgkdgntsaansadesvkgpnltelskkltdsnavllavkeveallissideilaakai	84	
Qy	62	GKKIRHNNLDITFEYHNHNSGLAGAVAIPTLIKOKLDGLKNEGLKEKIDAAKCSFTFTNK	121	
Db	85	gkklngmgldceyhngsllegayaisllqklidgklneglkeidaackcselftnk	144	
Qy	122	LKEKHTDLGKEGVTADAKAAILKTNGTTKGAELGKLFESVEVLSKAAKEMLANSVKE	181	
Db	145	lkehldlidgevtdadakeaillktngtklgaeeigklfesvevlskaakemlansvke	204	
Qy	182	LTSPPVAES---PAMGNSNGKGDASTPAPADESAKGPMLTETSKRTTDSNAFVLAVKEV	238	
Db	205	ltsppvaesppkpsmwnsgkdqnt-saansadesvgnpnltslkskltssnavvlavkev	263	
Qy	239	ETLVISIDELAKKAIGOKIDNNNGTAAALNNONGSLLAGAAYASTITTEKLSKLKNEEK	298	
Db	264	elllssideila-kalgkklkndsvlndeedhngslisgailsnlltkklsaisakdsgeik	322	
Qy	299	TEIARAKKCSSEFTNKLKSGHADLGKQDATTDDHAKAAILKTTHATTDKGAKKFKDLFESVE	358	
Db	323	aetlekakcseefatlkghecdlgegyvtdnakkailktmdcklgsadeleleklfesvk	382	
Qy	359	GLIKAAQVALTNSVKEL 375		
Db	383	nlskaakemltnsvkel 399		
RESULT 12				
AAAB62713				
ID	AAAB62713	standard	Protein	377 AA
XX	AAAB62713			

Db 299 kaelekakcseeflaklgentldgkvgvtdnnaekailktnndktkgadelekifeav 358
QY 358 EGLKRAQVALTNSVKEL 375
Db 359 knlskaekemltnsykel 376

RESULT 14

AAB62725 standard; Protein; 378 AA.

AC AAB62725;
DT 03-APR-2001 (first entry)
DE Borrelia sp chimeric ospc protein SEQ ID NO: 54.
XX
XX Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
XX Chimeric - Borrelia sp.
XX Chimeric - Borrelia sp.
XX WO200078966-A1.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

18-JUN-1999; 99US-0140042.

(UNYV) UNIV NEW YORK STATE RES FOUND.
(BROO-) BROOK BIOTECHNOLOGIES INC.

Dattwyler RJ, Seinoz G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI; 2001-050113/06.

N-PSDB; AAF29029.

Compositions of ospc polypeptides from strains of Borrelia which cause
Lyme disease are used to immunize animals and detect immune responses
to Lyme disease -

Claim 43; Page 114-115; 160pp; English.

The present invention provides compositions comprising ospc proteins and
chimeric ospc proteins from members of the Borrelia genus. These may be
Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
vaccines against Borrelia infection, which is spread by ticks and leads
to Lyme disease.

Sequence 378 AA;

Query Match 52.8%; Score 1452; DB 22; Length 378;

Best Local Similarity 58.7%; Pred. No. 4; 1e-89;

Matches 330; Conservative 9; Mismatches 31; Indels 192; Gaps 4;

QY 1 MACNSGKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDETIARA 60
DB 1 macnsgkdgntsansadsvkgnltetstktidsnavllavkevallssidel-aka 59
QY 61 IGGKTHQNGCLTPEYVNHNSLAGAVAITLTOKRIDGK-NEGKLEKIDAKKCEETPT 119
DB 60 Iggkthngcltpeyvnhsllagavaitltokrlidgk-negklekidaakcseets 119
QY 120 NKLKKEHTDLGKEGYTDADAEAILKTNGT-KTKGAEELGKLFESYEVLSKAKEWLANS 178
DB 120 tkkknhdgldgkgytdadaeailktngt-ktkgaeelgklfesyevlskaakemlans 179
QY 179 VKELTSPVVAESPAMGSGNSGKGDASATNPADESAAGPNLTETSKITDSNAFLAVKEY 238
DB 180 vkeltspvv----- 188

QY 239 ETLVLSTDELAKKAIQKIDNNNGLAALNNQNSLAGAVAITLTETKSLKLNLELK 298
DB 189 ----- 188
QY 299 TEIAKAKCSEEFNTKLGSHADLGKQDATDHAKAAILKTHATDGAKEFPLESEV 358
DB 189 ----- 188
QY 359 GLIKAAQVALTNSVKELGHRNNSGDSASTNPDESAKGNPLTVYSKKTIDSNAFLAVKE 418
DB 189 ----- -hgnsygdasastnpdesakgnpltvyskktidsnaflavke 230
QY 419 VEALLSSIDELSKAIGKTKTNNOSTLNEANRNSLAGAVEISKLITOKLSTVNSGEIK 478
DB 231 veallssideiskaigtktkndgclneanrnsllagaveislitqklsvinseelk 290
QY 479 KIEKADCSQKFTTKLSDSHAEIGIOSVODDNKAAILKTHGTRDKAGKELEELFKSLES 538
DB 291 kikeadcsqkfttklksdshaeligisvqddnakkailkthgtrdkagakeleelfsles 350
QY 539 LSKAAQALTNSVKELTNPVA 560
DB 351 lskaagaltnsykeltnpva 372

RESULT 15

AAB62733 standard; Protein; 401 AA.

AAB62733;

DT 03-APR-2001 (first entry)

Borrelia sp chimeric ospc protein SEQ ID NO: 70.

Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

Chimeric - Borrelia sp.
Chimeric - Borrelia sp.

WO200078966-A1.

28-DEC-2000.

19-JUN-2000; 2000WO-US16915.

18-JUN-1999; 99US-0140042.

(UNYV) UNIV NEW YORK STATE RES FOUND.
(BROO-) BROOK BIOTECHNOLOGIES INC.

Dattwyler RJ, Seinoz G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

WPI; 2001-050113/06.

N-PSDB; AAF29037.

Compositions of ospc polypeptides from strains of Borrelia which cause
Lyme disease are used to immunize animals and detect immune responses
to Lyme disease -

Claim 43; Page 136-137; 160pp; English.

The present invention provides compositions comprising ospc proteins and
chimeric ospc proteins from members of the Borrelia genus. These may be
Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
vaccines against Borrelia infection, which is spread by ticks and leads
to Lyme disease.

Sequence 401 AA;

Query Match 52.5%; Score 1444; DB 22; Length 401;
Best Local Similarity 58.5%; Pred. No. 1; 5e-88;

Matches 328; Conservative 10; Mismatches 31; Indels 192; Gaps 4;

```

QY      2 ACNNSGKDNTSANSADSVKGNPLTEISKRTDTSNAVLAVKEVEALLSSIDEIAAKAI 61
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     25 scnmsgkgnscansadesvkgpnltelskktltsdenavllavkeveallssidel-akai 83
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      62 GKTIHQNNGLDTEYNHNSLLAGAYAI STLKQKIDGLK-NEGLEKRTDAKKCSEFTN 120
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      84 gkklkndgslidneanrnesllagayl stlltqkrlskingseglkelaakkcseefst 143
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     121 KLKKEHTDLGEGYTDADAKFAILLKTNGT-KTKGAEEGLKLFESYEVL SKAAKEMLANSV 179
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     144 klkdnhaqlgqyvcdenakkaillkanaagdkyveeleklsgslskskaakemlansv 203
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     180 KELTSPVVAESPAMGNSNGSGDSASTNPADESA KGNPLTEISKRTDTSNAFVLAVKEVE 239
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     204 keltsprv----- 211
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     240 TLVLSIDELAKKAIGQKIDNNNGLA LNNQNGSLAGAYAI STLTERKLSKLNLEELKT 299
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     212 ----- 211
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     300 EIAKAKKCEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFKDLFESYEG 359
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     212 ----- 211
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     360 LLKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPNLTVISKRTDTSNAFLLAVKEV 419
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     212 ----- -hgnsngdsastnpdesakgnltvlskktltsnaflavkev 254
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     420 EALLSSIDELSKAIGKIKNDGTLDNENANRNE SLIAGAYEISKLITQKLSVLNSEEELKK 479
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     255 eallssidelaklqkklkndgtlidsneanrnesllagayelsklltqklsvlnseelkkk 314
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     480 IKEAKDCSQKFTTKLKDASHALGIOSVODDNAAKKA I LKTHGTDKGAKELEELFKLSLESL 539
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     315 lkeakdcsqkfttklkdsahaelgiqsvgd dnakkaillkthgtkdakeleelfkxlesl 374
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     540 SKAAQALTN SVKELTNPVVA 560
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     375 skaagaa ltnsvkeltnpvva 395
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: March 18, 2002, 09:54:35
 Job time: 333 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:08:46 ; Search time 621.2 Seconds
(without alignments)
250.303 Million cell updates/sec

Title: US-09-596-746A-52
Perfect score: 2750
Sequence: 1 MACNNSCKDNTSANSADSE.....KAAQALTNKYKELTNPVYA 560

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 27657034 residues
al number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US09_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2750	100.0	560	19	US-09-596-746-52
2	2750	100.0	560	19	US-09-596-746A-52
3	1758	63.9	410	19	US-09-596-746-84
4	1758	63.9	411	19	US-09-596-746A-84
5	1754	63.8	386	19	US-09-596-746-58
6	1754	63.8	386	19	US-09-596-746A-58
7	1734.5	63.1	385	19	US-09-596-746A-56
8	1729.5	62.9	384	19	US-09-596-746-56
9	1726.5	62.8	408	19	US-09-596-746-78

	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	1726.5	1496.5	1492.5	1491.5	1488.5	1488.5	1487.5	1484.5	1484.5	1474.5	1474.5	1469.5	1469.5	1465.5	1465.5	1460.5	1460.5	1452.5	1452.5	1444.5	1444.5	1203.5	1203.5	1199.5	1199.5	1198.5	1198.5	1195.5	1195.5	1195.5	1195.5	1195.5	1195.5	1167.5	1167.5	1167.5
	62.8	54.4	54.3	54.2	54.1	54.1	54.1	54.0	54.0	53.6	53.6	53.4	53.4	53.3	53.3	53.1	53.1	52.8	52.8	52.5	52.5	43.8	43.8	43.6	43.6	43.6	43.6	43.5	43.5	43.5	43.5	43.5	42.5	42.5	42.5	
	409	375	374	373	397	398	374	398	399	400	401	402	402	377	377	378	378	378	378	401	402	370	370	371	371	369	369	393	393	393	394	370	370	368	368	
	US-09-596-746A-78	US-09-596-746A-26	US-09-596-746A-24	US-09-596-746-26	US-09-596-746-62	US-09-596-746A-62	US-09-596-746-24	US-09-596-746-60	US-09-596-746A-60	US-09-596-746-82	US-09-596-746A-82	US-09-596-746-80	US-09-596-746A-80	US-09-596-746-30	US-09-596-746A-30	US-09-596-746-28	US-09-596-746A-28	US-09-596-746-54	US-09-596-746A-54	US-09-596-746-70	US-09-596-746A-70	US-09-596-746A-32	US-09-596-746A-32	US-09-596-746A-36	US-09-596-746A-34	US-09-596-746-32	US-09-596-746-36	US-09-596-746-66	US-09-596-746-66	US-09-596-746-68	US-09-596-746A-68	US-09-596-746A-68	US-09-596-746-64	US-09-596-746A-64	US-09-596-746-38	US-09-596-746A-38
	Sequence 78, Appl	Sequence 26, Appl	Sequence 26, Appl	Sequence 26, Appl	Sequence 62, Appl	Sequence 62, Appl	Sequence 24, Appl	Sequence 60, Appl	Sequence 60, Appl	Sequence 82, Appl	Sequence 82, Appl	Sequence 80, Appl	Sequence 80, Appl	Sequence 30, Appl	Sequence 30, Appl	Sequence 28, Appl	Sequence 28, Appl	Sequence 54, Appl	Sequence 54, Appl	Sequence 70, Appl	Sequence 70, Appl	Sequence 32, Appl	Sequence 32, Appl	Sequence 36, Appl	Sequence 34, Appl	Sequence 32, Appl	Sequence 36, Appl	Sequence 66, Appl	Sequence 66, Appl	Sequence 68, Appl	Sequence 68, Appl	Sequence 64, Appl	Sequence 64, Appl	Sequence 38, Appl	Sequence 38, Appl	

ALIGNMENTS

RESULT 1
US-09-596-746-52
Sequence 52, Application US/09596746
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seino, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
FILE REFERENCE: 2631.1002-001
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 52
LENGTH: 560
TYPE: PRT
ORGANISM: ospc Chimera
US-09-596-746-52

Query Match 100.0%; Score 2750; DB 19; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.1e-180;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACNNSCKDNTSANSADSEYKGNLTETISKKITDSNAVLAVKEVALLSIDEIATAKA 60
DB 1 MACNNSCKDNTSANSADSEYKGNLTETISKKITDSNAVLAVKEVALLSIDEIATAKA 60

```
QY 61 IGGKIHONNGDTEYNHNSLAGAVALSTLIKOKLDGKNEGLKEKIDAAKCSSEFTN 120
Db 61 IGGKIHONNGDTEYNHNSLAGAVALSTLIKOKLDGKNEGLKEKIDAAKCSSEFTN 120
QY 121 KLEKHTDGLGEGVTDADAKAAILKTNGTKGAELGKLFESVEVLSKAAKEMLANSVK 180
Db 121 KLEKHTDGLGEGVTDADAKAAILKTNGTKGAELGKLFESVEVLSKAAKEMLANSVK 180
QY 181 ELTSPVAESPAMSGNSGKGDASSTNPDESAGPNLTETSKITDSNAFVLAKEVEET 240
Db 181 ELTSPVAESPAMSGNSGKGDASSTNPDESAGPNLTETSKITDSNAFVLAKEVEET 240
QY 241 LVLSIDELAKKAIGOKIDNNNGLAALNNONGSLAGAVASTLITTEKLSKLNLEELKTE 300
Db 241 LVLSIDELAKKAIGOKIDNNNGLAALNNONGSLAGAVASTLITTEKLSKLNLEELKTE 300
QY 301 IAKAKCSEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
Db 301 IAKAKCSEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
QY 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPNLTETSKITDSNAFVLAKEVE 420
Db 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPNLTETSKITDSNAFVLAKEVE 420
QY 421 ALLSIDELSKAIGKIKINDGTLIDNEANRNESLAGAVEISKLITOKLSVNSEELKKI 480
Db 421 ALLSIDELSKAIGKIKINDGTLIDNEANRNESLAGAVEISKLITOKLSVNSEELKKI 480
QY 481 KEAKCSOKFTTKLKDISHAELGIGSVODDAAKKAAILKTHGTRKGAKELEELFKSLESLS 540
Db 481 KEAKCSOKFTTKLKDISHAELGIGSVODDAAKKAAILKTHGTRKGAKELEELFKSLESLS 540
QY 541 KAAQAAALTNSVKELTNPVVA 560
Db 541 KAAQAAALTNSVKELTNPVVA 560

RESULT 2
US-09-596-746a-52
; Sequence 52, Application US/09596746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 560
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746a-52
```

Query Match 100.0%; Score 2750; DB 19; Length 560;
Best Local Similarity 100.0%; Pred. No. 1,1e-180;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MACNNSGKDGNTSANSADSVKGNLTETSKITDSNAVLLAVEVEALLSSIDEIAKA 60
Db 1 MACNNSGKDGNTSANSADSVKGNLTETSKITDSNAVLLAVEVEALLSSIDEIAKA 60
QY 61 IGGKIHONNGDTEYNHNSLAGAVALSTLIKOKLDGKNEGLKEKIDAAKCSSEFTN 120
Db 61 IGGKIHONNGDTEYNHNSLAGAVALSTLIKOKLDGKNEGLKEKIDAAKCSSEFTN 120
```

```
Db 61 IGGKIHONNGDTEYNHNSLAGAVALSTLIKOKLDGKNEGLKEKIDAAKCSSEFTN 120
QY 121 KLEKHTDGLGEGVTDADAKAAILKTNGTKGAELGKLFESVEVLSKAAKEMLANSVK 180
Db 121 KLEKHTDGLGEGVTDADAKAAILKTNGTKGAELGKLFESVEVLSKAAKEMLANSVK 180
QY 181 ELTSPVAESPAMSGNSGKGDASSTNPDESAGPNLTETSKITDSNAFVLAKEVEET 240
Db 181 ELTSPVAESPAMSGNSGKGDASSTNPDESAGPNLTETSKITDSNAFVLAKEVEET 240
QY 241 LVLSIDELAKKAIGOKIDNNNGLAALNNONGSLAGAVASTLITTEKLSKLNLEELKTE 300
Db 241 LVLSIDELAKKAIGOKIDNNNGLAALNNONGSLAGAVASTLITTEKLSKLNLEELKTE 300
QY 301 IAKAKCSEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
Db 301 IAKAKCSEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGL 360
QY 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPNLTETSKITDSNAFVLAKEVE 420
Db 361 LKAAQVALTNSVKELGHRNNSGDSASTNPDESAGPNLTETSKITDSNAFVLAKEVE 420
QY 421 ALLSIDELSKAIGKIKINDGTLIDNEANRNESLAGAVEISKLITOKLSVNSEELKKI 480
Db 421 ALLSIDELSKAIGKIKINDGTLIDNEANRNESLAGAVEISKLITOKLSVNSEELKKI 480
QY 481 KEAKCSOKFTTKLKDISHAELGIGSVODDAAKKAAILKTHGTRKGAKELEELFKSLESLS 540
Db 481 KEAKCSOKFTTKLKDISHAELGIGSVODDAAKKAAILKTHGTRKGAKELEELFKSLESLS 540
QY 541 KAAQAAALTNSVKELTNPVVA 560
Db 541 KAAQAAALTNSVKELTNPVVA 560

RESULT 3
US-09-596-746-84
; Sequence 84, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selmost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 410
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-84
```

Query Match 63.9%; Score 1758; DB 19; Length 410;
Best Local Similarity 95.1%; Pred. No. 1,6e-112;
Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

```
QY 188 AEPAMSGNSGKGDASSTNPDESAGPNLTETSKITDSNAFVLAKEVEETLVLSIDE 247
Db 20 AEPISGNSGKGDASSTNPDESAGPNLTETSKITDSNAFVLAKEVEETLVLSIDE 79
QY 248 LAKKAIGOKIDNNNGLAALNNONGSLAGAVASTLITTEKLSKLNLEELKEIKAKKC 307
Db 80 LAKKAIGOKIDNNNGLAALNNONGSLAGAVASTLITTEKLSKLNLEELKEIKAKKC 139
QY 308 SEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFDLFESEVGLKAAQVA 367
```



```

|||||
Db 140 SEEFTHKLSGSHADLGKODATDHAHAAALTKHTATTDKGAKEKDFESEVEGLLKAQVA 199
OY 368 LTNVKEVLGH-----RNNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 415
Db 200 LTNVKEVLTPSPVAESPCKPHMANNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 259
OY 416 VKEVEALLSIDELSKAIGKIKKINDGTLDNEANRNESLAGAVEISKLITOKLSVANSSE 475
Db 260 VKEVEALLSIDELSKAIGKIKKINDGTLDNEANRNESLAGAVEISKLITOKLSVANSSE 319
OY 476 LKKIKKADCSOKFTTKLKDASHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 535
Db 320 LKKIKKADCSOKFTTKLKDASHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 379
OY 536 LESLSKAAQALTNVSKELTNPVYA 560
Db 380 LESLSKAAQALTNVSKELTNPVYA 404

```

JULT 4
 US-596-746A-84

Sequence 84, Application US/09596746A

GENERAL INFORMATION:

APPLICANT: Datwyler, Raymond J.

APPLICANT: Seimost, Gerald

APPLICANT: Dykhuisen, Daniai

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001

CURRENT FILING DATE: 2000-06-19

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 84

LENGTH: 411

TYPE: PRT

ORGANISM: ospC Chlamera

US-09-596-746A-84

Query Match 63.9%; Score 1758; DB 19; Length 411;

Best Local Similarity 95.1%; Pred. No. 1.6e-112;

Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

```

188 ASPPAMGNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEVLVL5IDE 247
Db 21 ASISGCSNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEVLVL5IDE 80
OY 248 LKKAIKGIKIDNNGLAALNNONGSLAGAVAIISLITEKLSKLNLEELKTEIAKAKC 307
Db 81 LKKAIKGIKIDNNGLAALNNONGSLAGAVAIISLITEKLSKLNLEELKTEIAKAKC 140
OY 308 SEEFTHKLSGSHADLGKODATDHAHAAALTKHTATTDKGAKEKDFESEVEGLLKAQVA 367
Db 141 SEEFTHKLSGSHADLGKODATDHAHAAALTKHTATTDKGAKEKDFESEVEGLLKAQVA 200
OY 368 LTNVKEVLGH-----RNNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 415
Db 201 LTNVKEVLTPSPVAESPCKPHMANNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLA 260
OY 416 VKEVEALLSIDELSKAIGKIKKINDGTLDNEANRNESLAGAVEISKLITOKLSVANSSE 475
Db 261 VKEVEALLSIDELSKAIGKIKKINDGTLDNEANRNESLAGAVEISKLITOKLSVANSSE 319
OY 476 LKKIKKADCSOKFTTKLKDASHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 535
Db 321 LKKIKKADCSOKFTTKLKDASHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKS 380

```

```

OY 536 LESLSKAAQALTNVSKELTNPVYA 560
Db 381 LESLSKAAQALTNVSKELTNPVYA 405

```

RESULT 5

US-09-596-746-58

Sequence 58, Application US/09596746

GENERAL INFORMATION:

APPLICANT: Datwyler, Raymond J.

APPLICANT: Seimost, Gerald

APPLICANT: Dykhuisen, Daniai

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001

CURRENT FILING DATE: 2000-06-16

PRIOR APPLICATION NUMBER: US 60/140,042

PRIOR FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 58

LENGTH: 386

TYPE: PRT

ORGANISM: ospC Chlamera

US-09-596-746-58

Query Match 63.8%; Score 1754; DB 19; Length 386;

Best Local Similarity 95.8%; Pred. No. 2.8e-112;

Matches 364; Conservative 0; Mismatches 4; Indels 12; Gaps 1;

```

OY 193 MGSNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEVLVL5IDE LAKKA 252
Db 1 MCSNSGKGGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVKEVEVLVL5IDE LAKKA 60
OY 253 IGGKIDNNNGLAALNNONGSLAGAVAIISLITEKLSKLNLEELKTEIAKAKCSEFT 312
Db 61 IGGKIDNNNGLAALNNONGSLAGAVAIISLITEKLSKLNLEELKTEIAKAKCSEFT 120
OY 313 NKLSGSHADLGKODATDHAHAAALTKHTATTDKGAKEKDFESEVEGLLKAQVALTNV 372
Db 121 NKLSGSHADLGKODATDHAHAAALTKHTATTDKGAKEKDFESEVEGLLKAQVALTNV 180
OY 373 KEIGH-----RNNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVEVE 420
Db 181 KEITSPVAESPCKPHMANNSGDSASTNPDESAGKPNLTVISKRTIDSNAPFLAVEVE 240
OY 421 ALLSIDELSKAIGKIKKINDGTLDNEANRNESLAGAVEISKLITOKLSVANSSEELKKI 480
Db 241 ALLSIDELSKAIGKIKKINDGTLDNEANRNESLAGAVEISKLITOKLSVANSSEELKKI 300
OY 481 KEAKDCSOKFTTKLKDASHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKSLESLS 540
Db 301 KEAKDCSOKFTTKLKDASHAELGIOSVODDNAAKKAAILKTHGTGKDGAKLEELFKSLESLS 360
OY 541 KAAQALTNVSKELTNPVYA 560
Db 361 KAAQALTNVSKELTNPVYA 380

```

RESULT 6

US-09-596-746A-58

Sequence 58, Application US/09596746A

GENERAL INFORMATION:

APPLICANT: Datwyler, Raymond J.

APPLICANT: Seimost, Gerald

APPLICANT: Dykhuisen, Daniai

APPLICANT: Luft, Benjamin J.

APPLICANT: Maria J.C. Gomes-Solecki

TITLE OF INVENTION: Groups of Borrelia burgdorferi and

;; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;; FILE REFERENCE: 2631.1002-001
;; CURRENT APPLICATION NUMBER: US/09/596,746A
;; CURRENT FILING DATE: 2000-06-19
;; PRIOR APPLICATION NUMBER: US 60/140,042
;; PRIOR FILING DATE: 1999-06-18
;; NUMBER OF SEQ ID NOS: 84
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 58
;; LENGTH: 386
;; TYPE: PRF
;; ORGANISM: ospc Chimera
US-09-596-746a-58

Query Match 63.8%; Score 1754; DB 19; Length 386;
Best Local Similarity 95.8%; Pred. No. 2.8e-112;
Matches 364; Conservative 0; Mismatches 4; Indels 12; Gaps 1;

193 MGSNSGSGSASTNPADBSAKGNLTETSKITTSNAFLAVKEVEVLVLSIDELAKKA 252
1 MGSNSGSGSASTNPADBSAKGNLTETSKITTSNAFLAVKEVEVLVLSIDELAKKA 60
253 IGGKIDNNGLAALNNONGSLAGAYAISTLTETKLSKLNLELTETETAKAKCSEFT 312
61 IGGKIDNNGLAALNNONGSLAGAYAISTLTETKLSKLNLELTETETAKAKCSEFT 120
313 NKLSGSHADIGKODATDHAKAAILKTHATTGKAKEFDFESVGLKKAQVALTNSV 372
121 NKLSGSHADIGKODATDHAKAAILKTHATTGKAKEFDFESVGLKKAQVALTNSV 180
373 KELGH-----RNSGSGDSASTNPDSAKGNLTETSKITTSNAFLAVKEVE 420
181 KELTSVVAESPMPKPMANNSGDSASTNPDSAKGNLTETSKITTSNAFLAVKEVE 240
421 ALLSIDELSKAIGKIKNDGTLDNEANRNESLIGAYEISLTOKLSVNSEELKKI 480
241 ALLSIDELSKAIGKIKNDGTLDNEANRNESLIGAYEISLTOKLSVNSEELKKI 300
481 KEAKDSQKFTTKLDSHAELGIQSVODNNAKKAAILKTHGTGKAKELEELFKSLESLS 540
301 KEAKDSQKFTTKLDSHAELGIQSVODNNAKKAAILKTHGTGKAKELEELFKSLESLS 360
541 KAAQALTNVSKELTNPVVA 560
361 KAAQALTNVSKELTNPVVA 380

BLT 7
US-09-596-746a-56
;; Sequence 56, Application US/09596746A
;; GENERAL INFORMATION:
;; APPLICANT: Dattwyler, Raymond J.
;; APPLICANT: Seinoest, Gerald
;; APPLICANT: Dykhuisen, Daniel
;; APPLICANT: Luft, Benjamin J.
;; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;; FILE REFERENCE: 2631.1002-001
;; CURRENT APPLICATION NUMBER: US/09/596,746A
;; CURRENT FILING DATE: 2000-06-19
;; PRIOR APPLICATION NUMBER: US 60/140,042
;; PRIOR FILING DATE: 1999-06-18
;; NUMBER OF SEQ ID NOS: 84
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 56
;; LENGTH: 385
;; TYPE: PRF
;; ORGANISM: ospc Chimera
US-09-596-746a-56

Query Match 63.1%; Score 1734.5; DB 19; Length 385;
Best Local Similarity 67.0%; Pred. No. 6.1e-111;
Matches 375; Conservative 1; Mismatches 3; Indels 181; Gaps 1;

1 MACNNSGKGNFSANSADSVGPNLTETSKITTSNAFLAVKEVEALLSIDEIATAKA 60
1 MACNNSGKGNFSANSADSVGPNLTETSKITTSNAFLAVKEVEALLSIDEIATAKA 60
61 ICKKHONNGLDTEYNNHNSLAGAYAISTLTOKLDGKNGEKIDAAKCSSEFTN 120
61 ICKKHONNGLDTEYNNHNSLAGAYAISTLTOKLDGKNGEKIDAAKCSSEFTN 120
121 KKEKHTDGLRGVYDADAKEAILKTNGTKTGABELKLEFESVGLSKAAEMLANSVK 180
121 KKEKHTDGLRGVYDADAKEAILKTNGTKTGABELKLEFESVGLSKAAEMLANSVK 180
181 ELTSPVAESPAMGNSGSGDSASTNPADBSAKGNLTETSKITTSNAFLAVKEVE 240
181 ELTSPVAESP----- 191
241 LVLSIDELAKKAIGKIDNNGLAALNNONGSLAGAYAISTLTETKLSKLNLEEKTE 300
192 ----- 191
301 IAKAKCSEFTNKLKSGHADIGKODATDHAKAAILKTHATTGKAKERDLPESVEGL 360
192 ----- 191
361 LKAQVALTNVSKELGHRNNSGDSASTNPDSAKGNLTETSKITTSNAFLAVKEVE 420
192 -----KPFHGNNSGDSASTNPDSAKGNLTETSKITTSNAFLAVKEVE 239
421 ALLSIDELSKAIGKIKNDGTLDNEANRNESLIGAYEISLTOKLSVNSEELKKI 480
240 ALLSIDELSKAIGKIKNDGTLDNEANRNESLIGAYEISLTOKLSVNSEELKKI 299
481 KEAKDSQKFTTKLDSHAELGIQSVODNNAKKAAILKTHGTGKAKELEELFKSLESLS 540
300 KEAKDSQKFTTKLDSHAELGIQSVODNNAKKAAILKTHGTGKAKELEELFKSLESLS 359
541 KAAQALTNVSKELTNPVVA 560
360 KAAQALTNVSKELTNPVVA 379

RESULT 8
US-09-596-746-56
;; Sequence 56, Application US/09596746
;; GENERAL INFORMATION:
;; APPLICANT: Dattwyler, Raymond J.
;; APPLICANT: Seinoest, Gerald
;; APPLICANT: Dykhuisen, Daniel
;; APPLICANT: Luft, Benjamin J.
;; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;; FILE REFERENCE: 2631.1002-001
;; CURRENT APPLICATION NUMBER: US/09/596,746
;; CURRENT FILING DATE: 2000-06-16
;; PRIOR APPLICATION NUMBER: US 60/140,042
;; PRIOR FILING DATE: 1999-06-18
;; NUMBER OF SEQ ID NOS: 84
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 56
;; LENGTH: 384
;; TYPE: PRF
;; ORGANISM: ospc Chimera
US-09-596-746-56

Query Match 62.9%; Score 1729.5; DB 19; Length 384;
Best Local Similarity 66.9%; Pred. No. 1.3e-110;
Matches 374; Conservative 1; Mismatches 3; Indels 181; Gaps 1;

```
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 1 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 60
QY 62 GKRIHONNGDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 121
DB 61 GKRIHONNGDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 120
QY 122 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 121 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 180
QY 182 LTSPVVAESPAMSGSGKGDASTNPDESAGKPNLITEISKITDSNAVLAVKEVETL 241
DB 181 LTSPVVAESP----- 190
QY 242 VLSIDELAKKAIGOKIDNNGGLAALNNGSLLAGAYAI STLTEKLSKLNLELKEI 301
DB 191 ----- 190
QY 302 AKAKKCEFTNKLKSGHADLGKODATDHAKAAILKTHATTDKGAKEFPDLPESVEGL 361
DB 191 ----- 190
QY 362 KAAQVALTNSVKELGHRNNSGDSASTNPDESAGKPNLTVYISKITDSNAFLAVKEVEA 421
DB 191 ----- 190
QY 191 -----KRFHGNNSGDSASTNPDESAGKPNLTVYISKITDSNAFLAVKEVEA 239
QY 422 LLSIDELSAIKGKIKNDGTLDEANRNESLAGAYEISKLITOKLSVNSEELKRIK 481
DB 240 LLSIDELSAIKGKIKNDGTLDEANRNESLAGAYEISKLITOKLSVNSEELKRIK 299
QY 482 EAKDCSOKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 541
DB 300 EAKDCSOKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 359
QY 542 AAQALTNVSKELTNPYVA 560
DB 360 AAQALTNVSKELTNPYVA 378

RESULT 9
US-09-596-746-78
; Sequence 78, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 408
; TYPE: PRN
; ORGANISM: ospC Chimera
US-09-596-746-78

Query Match 62.8%; Score 1726.5; DB 19; Length 408;
Best Local Similarity 66.7%; Pred. No. 2.4e-110;
Matches 373; Conservative 2; Mismatches 3; Indels 181; Gaps 1;
```

```
DB 25 SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 84
QY 62 GKRIHONNGDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 121
DB 85 GKRIHONNGDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 144
QY 122 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 181
DB 145 LKEKHTDLGEGVTDADAKAEAILKTNGTKTGAEELGKLFESVEVLSKAKEMLANSVKE 204
QY 182 LTSPVVAESPAMSGSGKGDASTNPDESAGKPNLITEISKITDSNAVLAVKEVETL 241
DB 205 LTSPVVAESP----- 214
QY 242 VLSIDELAKKAIGOKIDNNGGLAALNNGSLLAGAYAI STLTEKLSKLNLELKEI 301
DB 215 ----- 214
QY 302 AKAKKCEFTNKLKSGHADLGKODATDHAKAAILKTHATTDKGAKEFPDLPESVEGL 361
DB 215 ----- 214
QY 362 KAAQVALTNSVKELGHRNNSGDSASTNPDESAGKPNLTVYISKITDSNAFLAVKEVEA 421
DB 215 -----KRFHGNNSGDSASTNPDESAGKPNLTVYISKITDSNAFLAVKEVEA 263
QY 422 LLSIDELSAIKGKIKNDGTLDEANRNESLAGAYEISKLITOKLSVNSEELKRIK 481
DB 264 LLSIDELSAIKGKIKNDGTLDEANRNESLAGAYEISKLITOKLSVNSEELKRIK 323
QY 482 EAKDCSOKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 541
DB 324 EAKDCSOKFTTKLDSHAELGISOVDNNAKKAILKTHGTRKDGAKLEELFKLSLESLK 383
QY 542 AAQALTNVSKELTNPYVA 560
DB 384 AAQALTNVSKELTNPYVA 402

RESULT 10
US-09-596-746a-78
; Sequence 78, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuzen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 409
; TYPE: PRN
; ORGANISM: ospC Chimera
US-09-596-746a-78

Query Match 62.8%; Score 1726.5; DB 19; Length 409;
Best Local Similarity 66.7%; Pred. No. 2.4e-110;
Matches 373; Conservative 2; Mismatches 3; Indels 181; Gaps 1;
```

```
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 62 GKRIHONNGDTEYNHNGSLLAGAYAI STLTKOKLDGLKNEGKLEKIDAAKCCSEFTFNK 121
```

```
QY 2 ACNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 61
DB 26 SCNNSGKDGNTSANSADSVKGPMLTEISKITDSNAVLAVKEVEALLSSIDEIAAKAI 85
```

```

Db      86 GKKIHONNGIDTEYNHNSGLLAGAVAI STLKQKIDGLKNEGKEKIDAAKCSSEFTTK 145
Qy      122 LKEKHTDLGEGVTDADAKAEALIKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 181
Db      146 LKEKHTDLGEGVTDADAKAEALIKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 205
Qy      182 LTPSPVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAVKEVETL 241
Db      206 LTPSPVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAVKEVETL 215
Qy      242 VLSIDELAKKAIGKIDNNNGLAALNNONGSLAGAVAI STLTEKLSKLNLEELKTEI 301
Db      216 ----- 215
Qy      302 AKAKKCSSEFTTKLSGSHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 361
Db      216 ----- 215
Qy      362 KAAQVALTNSVKELGHRNNSGSDSASTNPDESAGPMLTVISKKITDSNAFVLAVKEVETL 421
Db      216 ----- 215
Qy      422 LLSIDELSKAIGKIKKNGTIDNENANRESLIGAVYEISKLITQKLSVNSEELKKKIK 481
Db      265 LLSIDELSKAIGKIKKNGTIDNENANRESLIGAVYEISKLITQKLSVNSEELKKKIK 324
Qy      482 EAKDCSQKFTTKIKDSHAELGIGSVODDANAKKAILKTHGTGKGAKELEFKLSLESLSK 541
Db      325 EAKDCSQKFTTKIKDSHAELGIGSVODDANAKKAILKTHGTGKGAKELEFKLSLESLSK 384
Qy      542 AAQVALTNSVKELTNPVVA 560
Db      385 AAQVALTNSVKELTNPVVA 403

RESULT 11
US-09-596-746a-26
; Sequence 26, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luit, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-26

```

```

Query Match      54.4%; Score 1496.5; DB 19; Length 374;
Best Local Similarity 82.7%; Pred. No. 1.5e-94;
Matches 310; Conservative 21; Mismatches 43; Indels 1; Gaps 1;

```

```

Qy      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAVKEVALLSSIDEIAKA 60
Db      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAVKEVALLSSIDEIAKA 60
Qy      61 IGKRIHONNGIDTEYNHNSGLLAGAVAI STLKQKIDGLKNEGKEKIDAAKCSSEFTTN 120

```

```

Db      61 IGKRIHONNGIDTEYNHNSGLLAGAVAI STLKQKIDGLKNEGKEKIDAAKCSSEFTTN 120
Qy      121 KLEKHTDLGEGVTDADAKAEALIKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
Db      121 KLEKHTDLGEGVTDADAKAEALIKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
Qy      181 ELTSPVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAVKEVETL 240
Db      181 ELTSPVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAVKEVETL 239
Qy      241 VLSIDELAKKAIGKIDNNNGLAALNNONGSLAGAVAI STLTEKLSKLNLEELKTE 300
Db      240 LLSIDELAKKAIGKIKKNGTIDNENANRESLIGAVYEISKLITQKLSVNSEELKKKIK 299
Qy      301 IAKAKKCSSEFTTKLSGSHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360
Db      300 IMAKKSSEFTTKLEGEHAQIGIENVTDENAKKAILITDAKKGAELEKFRVAVENL 359
Qy      361 LKAQVALTNSVKEL 375
Db      360 AKAKEMLANSVKEL 374

RESULT 12
US-09-596-746a-24
; Sequence 24, Application US/09596746a
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinoest, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luit, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746a
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746a-24

```

```

Query Match      54.3%; Score 1492; DB 19; Length 375;
Best Local Similarity 83.5%; Pred. No. 3e-94;
Matches 313; Conservative 20; Mismatches 40; Indels 2; Gaps 2;

```

```

Qy      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAVKEVALLSSIDEIAKA 60
Db      1 MACNNSGKDGNTSANSADSVKGPMLTEISKKITDSNAFVLAVKEVALLSSIDEIAKA 60
Qy      61 IGKRIHONNGIDTEYNHNSGLLAGAVAI STLKQKIDGLKNEGKEKIDAAKCSSEFTTN 120
Db      61 IGKRIHONNGIDTEYNHNSGLLAGAVAI STLKQKIDGLKNEGKEKIDAAKCSSEFTTN 120
Qy      121 KLEKHTDLGEGVTDADAKAEALIKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
Db      121 KLEKHTDLGEGVTDADAKAEALIKTNGTKTGAEEGLKFESVEVLSKAKEMLANSVK 180
Qy      181 ELTSPVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAVKEVETL 240
Db      181 ELTSPVAESPAMGNSGKGDASTNPADESAAGPMLTEISKKITDSNAFVLAVKEVETL 239
Qy      241 VLSIDELAKKAIGKIDNNNGLAALNNONGSLAGAVAI STLTEKLSKLNLEELKTE 300
Db      240 LLSIDELAKKAIGKIKKNGTIDNENANRESLIGAVYEISKLITQKLSVNSEELKKKIK 298

```


SEQ ID NO 62
LENGTH: 398
TYPE: PRT
ORGANISM: ospc Chimera
US-09-596-746a-62

Query Match 54.1%; Score 1488.5; DB 19; Length 398;
Best Local Similarity 82.4%; Pred. No. 5.8e-94;
Matches 308; Conservative 22; Mismatches 43; Indels 1; Gaps 1;

QY 2 ACNNSGKDNTSANSADSVKGPMLTEISKITDSSNAVLAKEVEALLSSIDEIAAKAI 61
: |||||
Db 26 SCNNSGKDNTSANSADSVKGPMLTEISKITDSSNAVLAKEVEALLSSIDEIAAKAI 85
QY 62 GKTIHONGLTDEYNHNSLAGAIAISTLIKOKLDGLKNEGLKEKIDAAKCSFTFNK 121
: |||||
Db 86 GKTIHONGLTDEYNHNSLAGAIAISTLIKOKLDGLKNEGLKEKIDAAKCSFTFNK 145
122 LKEHTDGLKEGVTDADAKAAILKTNGTKTGAEELGKLFESVEYLSKAKEMLANSYKE 181
: |||||
Db 146 LKEHTDGLKEGVTDADAKAAILKTNGTKTGAEELGKLFESVEYLSKAKEMLANSYKE 205
QY 182 LTSPVAESPAMGNSGKGDASTNPADSAKGPMLTEISKITDSSNAVLAKEVEETL 241
: |||||
Db 206 LTSPVAESPAMGNSGKGDASTNPADSAKGPMLTEISKITDSSNAVLAKEVEETL 264
QY 242 VLSIDELAKKAIGOKIDNNGLAALNNGNSLAGAIAISTLITEKLSKLNLEIKTEI 301
: |||||
Db 265 LASTIDELATKAIGKIOONGLAIEAGHNGTLGAYTISKLTOKLDGLKNSKLEKEI 324
QY 302 AKAKKCEEFNTKLKSGHADLGKODATDDHAKAAILKTHATTDKAKEFKDLFESVEGIL 361
: |||||
Db 325 ENAKKCEDEFKRLKEGHAQGIENVTDENAKKAILLITDAKDKGAELKLFKAVENLA 384
QY 362 KAAQVALTNSYKEL 375
: |||||
Db 385 KAAKEMLANSYKEL 398

Search completed: March 18, 2002, 10:08:47
Job time: 975 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:35 ; Search time 55.5 Seconds
(without alignments)
227.060 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750

Sequence: 1 MACNNSGKGNSTANSANSADES.....KAAQALITNSVKELTPVVA 560

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	986.5	35.9	466	4	US-08-235-836C-110 Sequence 110, Appl
2	932	33.9	210	4	US-08-235-836C-30 Sequence 30, Appl
3	928	33.7	209	4	US-09-196-293-15 Sequence 15, Appl
4	928	33.7	210	1	US-08-158-353-3 Sequence 3, Appl
5	922	33.5	588	4	US-08-235-836C-122 Sequence 122, Appl
6	921	33.5	210	4	US-08-209-603E-15 Sequence 15, Appl
7	915	33.3	466	4	US-08-235-836C-107 Sequence 107, Appl
8	889	32.3	212	4	US-08-235-836C-34 Sequence 34, Appl
9	886	32.2	212	1	US-08-158-353-4 Sequence 4, Appl
10	882	32.1	207	4	US-08-235-836C-36 Sequence 36, Appl
11	858	31.2	212	4	US-09-196-293-11 Sequence 11, Appl
12	858	31.2	212	4	US-08-209-603E-11 Sequence 11, Appl
13	676	24.6	212	1	US-08-158-353-2 Sequence 2, Appl
14	661	24.0	212	1	US-08-031-295-2 Sequence 2, Appl
15	661	24.0	212	4	US-07-903-580-2 Sequence 2, Appl
16	616.5	22.4	209	4	US-08-235-836C-32 Sequence 32, Appl
17	188	6.8	2285	4	US-09-308-375-2 Sequence 2, Appl
18	187	6.8	1164	4	US-08-923-992A-2 Sequence 2, Appl
19	186	6.8	1164	4	US-08-923-992A-2 Sequence 2, Appl
20	184	6.7	1164	4	US-08-923-992A-10 Sequence 10, Appl
21	183.5	6.6	1588	5	PCT-US93-07261-11 Sequence 11, Appl
22	182.5	6.6	1630	5	PCT-US93-07261-16 Sequence 16, Appl
23	181.5	6.6	630	4	US-08-973-462-9 Sequence 9, Appl
24	181	6.5	688	3	US-09-141-047-8 Sequence 8, Appl
25	179	6.5	1128	4	US-08-923-992A-6 Sequence 6, Appl
26	177.5	6.5	1098	4	US-08-923-992A-8 Sequence 8, Appl
27	177.5	6.5	1196	1	US-08-144-121-4 Sequence 4, Appl

28	177.5	6.5	1196	2	US-08-735-893-4 Sequence 4, Appl
29	177	6.4	1388	2	US-08-685-576-1 Sequence 1, Appl
30	176.5	6.4	3248	1	US-08-353-700-1 Sequence 1, Appl
31	176.5	6.4	3248	5	PCT-US95-16216-1 Sequence 1, Appl
32	175.5	6.4	2482	1	US-08-328-254-6 Sequence 6, Appl
33	174	6.3	740	5	US-08-257-073-5 Sequence 5, Appl
34	173	6.3	1038	4	US-09-541-787-4 Sequence 4, Appl
35	171	6.2	1104	4	US-08-923-992A-4 Sequence 4, Appl
36	171	6.2	1388	2	US-08-685-576-4 Sequence 4, Appl
37	170.5	6.2	1151	3	US-08-840-006-6 Sequence 6, Appl
38	170.5	6.2	1200	3	US-08-840-006-5 Sequence 5, Appl
39	170.5	6.2	3111	2	US-08-460-309-4 Sequence 4, Appl
40	170.5	6.2	3111	2	US-08-125-077-4 Sequence 4, Appl
41	167	6.1	1147	3	US-08-470-260-5 Sequence 5, Appl
42	167	6.1	1147	3	US-08-471-491-5 Sequence 5, Appl
43	167	6.1	1147	4	US-08-466-662-5 Sequence 5, Appl
44	167	6.1	2101	1	US-08-195-487-4 Sequence 4, Appl
45	167	6.1	2101	5	PCT-US93-06160-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising Botreitia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110
Query Match 35.9%; Score 986.5; DB 4; Length 466;
Best Local Similarity 51.1%; Pred. No. 86-63;
Matches 256; Conservative 42; Mismatches 132; Indels 71; Gaps 17;
QY 1 MACNNSGKGNSTANSANSADESVKGNPLTETISKTTDSNANVLAVKEVALLSIDEITAARA 60

```

Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVLLAVKEVALLSSIDEIAKA 76
QY 61 IGGKTHONNGIDTEYNNHNSGLAGYAISTLIKOKLDGLKNEGIKERIDAKKCEFTTN 120
Db 77 IGGKTHONNGIDTEYNNHNSGLAGYAISTLIKOKLDGLKNEGIKERIDAKKCEFTTN 136
QY 121 KLEKHTDGLKEGYTDADAKEAIIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 180
Db 137 KLEKHTDGLKEGYTDADAKEAIIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 196
QY 181 ELTSPVVAESPAMGNSGKGDSDASTNPDESAGKPNLTVEISKRTIDSNVLLAVKEV 240
Db 197 ELTSPVVAESPAMGNSGKGDSDASTNPDESAGKPNLTVEISKRTIDSNVLLAVKEV 246
QY 241 LVLSTIDELAKKAIGOKIDNNNGSLAALNNNGSLAAGYAISTLIKOKLDGLKNEGI 300
Db 247 LVLSTIDELAKKAIGOKIDNNNGSLAALNNNGSLAAGYAISTLIKOKLDGLKNEGI 298
Db 301 -----TAKAKCSEE-FTNKLKSGHADLGKODATDHAKAALIKTHATDNGAKKEF 350
Db 299 GKTIVSKKVTSKDKSSTEKEFNEKGEVSEKIITRADGTR-----LETTGKSDSGKA 351
QY 351 KDLFEES--VGLLKAQVALTNSVKELGHRNNSGDSASTNPDESAGKPNLTVEISKRT 408
Db 352 KEVLKGYVLEGITLAEKTTLV--VKE-----GVVTLKSNISKSGE---VSVELNDTD 398
QY 409 SNAFLAVEVEALLSSIDEIS-KAIGKKIKN-----DGLTNEA--NNESLIGAGYEI 460
Db 399 SSA--ATKKTAMNDSTLTISADSKRTKDLVFLTDGTLTYOQNTAGTLEGSASEI 455
QY 461 SKLITOKLVLNSEEKKRIK 481
Db 456 K-----NLSELKNALK 466

RESULT 2
US-08-235-836C-30
; Sequence 30, Application US/08235836C
; Patent No. 6248362
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESS: BrocKhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-Apr-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 30:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-30

Query Match          33.9%; Score 932; DB 4; Length 210;
Best Local Similarity 98.4%; Pred. No. 2,1e-59;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVLLAVKEVALLSSIDEIAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVLLAVKEVALLSSIDEIAKA 76
QY 61 IGGKTHONNGIDTEYNNHNSGLAGYAISTLIKOKLDGLKNEGIKERIDAKKCEFTTN 120
Db 77 IGGKTHONNGIDTEYNNHNSGLAGYAISTLIKOKLDGLKNEGIKERIDAKKCEFTTN 136
QY 121 KLEKHTDGLKEGYTDADAKEAIIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 180
Db 137 KLEKHTDGLKEGYTDADAKEAIIKTNGTKGAEEIGKLFESVEVLSKAKEMLANSVK 196
QY 181 ELTSPVVAESP 191
Db 197 ELTSPVVAESP 207

RESULT 3
US-09-196-293-15
; Sequence 15, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; APPLICANT: Fuchs, Renate
; APPLICANT: Motz, Manfred
; APPLICANT: Soutscheck, Erwin
; APPLICANT: Wilske, Bettina
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrelia
; FILE REFERENCE: 738.001US2
; CURRENT APPLICATION NUMBER: US/09/196,293
; EARLIER FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1989-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; EARLIER FILING DATE: 1990-06-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match          33.7%; Score 928; DB 4; Length 209;
Best Local Similarity 98.4%; Pred. No. 4e-59;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVLLAVKEVALLSSIDEIAKA 60
Db 17 ISCNNSGKDGNTSANSADSVKGNPLTEISKRTIDSNVLLAVKEVALLSSIDEIAKA 76
QY 61 IGGKTHONNGIDTEYNNHNSGLAGYAISTLIKOKLDGLKNEGIKERIDAKKCEFTTN 120

```


Db 77 IGKTIHONNGDLENNHNSGLACAVAI STLTKKIDGLKNEGLKEKIDAAKCSFTFN 136
Qy 121 KLEKHTDLGREGVTDADAKKAI LTKNGTKTGAEELGKLFESVYLSAAKEMLANSVK 180
Db 137 KLEKHTDLGREGVTDADAKKAI LTKNGTKTGAEELGKLFESVYLSAAKEMLANSVK 196
Qy 181 ELTSPVAESP 191
Db 197 ELTSPVAESP 207

RESULT 4

US-08-158-353-3

; Sequence 3, Application US/08158353

; Patent No. 5620862

; GENERAL INFORMATION:

; APPLICANT: Padula, Steven J.

; TITLE OF INVENTION: Methods for Diagnosing Early Lyme

; TITLE OF INVENTION: Disease

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Millcila Drive

; CITY: Lexington

; STATE: MA

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/158.353

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Carroll, Alice O.

; REGISTRATION NUMBER: 33,542

; REFERENCE/DOCKET NUMBER: DCT93-05

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 210 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; 08-158-353-3

Query Match 33.7%; Score 928; DB 1; Length 210;

Best Local Similarity 98.4%; Pred. No. 4e-59;

Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MACNNSGDKGNTSANSADSVKGNLTLTKITDSNAVLAVKEVEALLSSIDEIAAKA 60
Db 17 ISCNNSGDKGNTSANSADSVKGNLTLTKITDSNAVLAVKEVEALLSSIDEIAAKA 76
Qy 61 IGKTIHONNGDLENNHNSGLACAVAI STLTKKIDGLKNEGLKEKIDAAKCSFTFN 120
Db 77 IGKTIHONNGDLENNHNSGLACAVAI STLTKKIDGLKNEGLKEKIDAAKCSFTFN 136
Qy 121 KLEKHTDLGREGVTDADAKKAI LTKNGTKTGAEELGKLFESVYLSAAKEMLANSVK 180
Db 137 KLEKHTDLGREGVTDADAKKAI LTKNGTKTGAEELGKLFESVYLSAAKEMLANSVK 196
Qy 181 ELTSPVAESP 191
Db 197 ELTSPVAESP 207

RESULT 5

US-08-235-836C-122

; Sequence 122, Application US/08235836C

; Patent No. 6248562

; GENERAL INFORMATION:

; APPLICANT: Dunn, John J.

; TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising

; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor

; NUMBER OF SEQUENCES: 144

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brookhaven National Laboratory

; STREET:

; CITY: Upton

; STATE: NY

; COUNTRY: USA

; ZIP: 11973

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/235.836C

; FILING DATE: 29-APR-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/148,191

; FILING DATE: 01-11-93

; ATTORNEY/AGENT INFORMATION:

; NAME: Bogosian, Margaret C.

; REGISTRATION NUMBER: 25,324

; REFERENCE/DOCKET NUMBER: BNL93-28A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 282-7338

; TELEFAX: (516) 282-3729

; INFORMATION FOR SEQ ID NO: 122:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 588 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-235-836C-122

Query Match 33.5%; Score 922; DB 4; Length 588;

Best Local Similarity 93.6%; Pred. No. 4.3e-58;

Matches 191; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

Qy 4 NNSGDKGNTSANSADSVKGNLTLTKITDSNAVLAVKEVEALLSSIDEIAAKA 63
Db 286 NNSGDKGNTSANSADSVKGNLTLTKITDSNAVLAVKEVEALLSSIDEIAAKA 345
Qy 64 KTIHONNGDLENNHNSGLACAVAI STLTKKIDGLKNEGLKEKIDAAKCSFTFN 123
Db 346 KTIHONNGDLENNHNSGLACAVAI STLTKKIDGLKNEGLKEKIDAAKCSFTFN 405
Qy 124 EKHTDLGREGVTDADAKKAI LTKNGTKTGAEELGKLFESVYLSAAKEMLANSVK 183
Db 406 EKHTDLGREGVTDADAKKAI LTKNGTKTGAEELGKLFESVYLSAAKEMLANSVK 465
Qy 184 SPVVAESP 207
Db 466 SPVVAESP 483

RESULT 6
US-08-209-603E-15
; Sequence 15, Application US/08209603E
; Patent No. 6248538
; GENERAL INFORMATION:
; APPLICANT: FUCHS, RENATE

APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURISIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: IIR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 557-3535
TELEFAX: (212) 557-3535
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

Query Match 33.5%; Score 921; DB 4; Length 210;
Best Local Similarity 97.9%; Pred. No. 1.3e-58;
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MACNSKDGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDEIAAKA 60
DB 17 TSCNNSGDKGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDEIAAKA 76
QY 61 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKOKLDGKNEGKKEKIDAKKCEFTFN 120
DB 77 IGGKIHONNGLDTEYNNHNSLLAGAYASTLIKOKLDGKNEGKKEKIDAKKCEFTFN 136
QY 121 KLEKHTDLGREGVTDADAKKAILKTNGTKGAEELGKLFESVEVLSKAKEMLANSYK 180
DB 137 KLEKHTDLGREGVTDADAKKAILKTNGTKGAEELGKLFESVEVLSKAKEMLANSYK 196
QY 181 ELTSPVAESP 191

DB 197 ELTSPVAESP 207
RESULT 7
US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
Title Of Invention: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-107

Query Match 33.3%; Score 915; DB 4; Length 466;
Best Local Similarity 98.9%; Pred. No. 1e-57;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 NNSGDKGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDEIAAKAIGK 63
DB 276 NNSGDKGNTSANSADSVKGNLTETSKITDSNAVLAVKEVALLSIDEIAAKAIGK 335
QY 64 KIHONNGLDTEYNNHNSLLAGAYASTLIKOKLDGKNEGKKEKIDAKKCEFTFN 123
DB 336 KIHONNGLDTEYNNHNSLLAGAYASTLIKOKLDGKNEGKKEKIDAKKCEFTFN 183
QY 124 EKHTDLGKGVTDADAKKAILKTNGTKGAEELGKLFESVEVLSKAKEMLANSYK 183
DB 396 AKHTDLGKGVTDADAKKAILKTNGTKGAEELGKLFESVEVLSKAKEMLANSYK 455
QY 184 SPVAESP 191
DB 456 SPVAESP 463
RESULT 8
US-08-235-836C-34
Sequence 34, Application US/08235836C

```

: Patent No.6248562
: GENERAL INFORMATION:
: APPLICANT: Dunn, John J.
: APPLICANT: Luft, Benjamin J.
: TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
: TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
: NUMBER OF SEQUENCES: 144
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brookhaven National Laboratory
: STREET:
: CITY: Upton
: STATE: NY
: COUNTRY: USA
: ZIP: 11973

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US//08/235,836C
: FILING DATE: 29-APR-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/148,191
: FILING DATE: 01-11-93
: ATTORNEY/AGENT INFORMATION:
: NAME: Bogosian, Margaret C.
: REGISTRATION NUMBER: 25,324
: REFERENCE/DOCKET NUMBER: BNL93-28A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 282-7338
: TELEFAX: (516) 282-3729
: INFORMATION FOR SEQ ID NO: 34:
: LENGTH: 212 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-235-836C-34

Query Match      32.3%, Score 889; DB 4; Length 212;
Best Local Similarity 91.2%; Pred. No. 2.4e-56;
Matches 187; Conservative 1; Mismatches 3; Indels 14; Gaps 2.

QY 195 SNSGCGGSASINPNPDESAKGNLTREISKRTIDSNAPFLVAKEVETLYSIDELAKKAIG 254
    |||||||
DB 20 SNSGCGGSASINPNPDESAKGNLTREISKRTIDSNAPFLVAKEVETLYSIDELAKKAIG 79
    |||||||
255 OKIDNNNGILALNNGSILAGAVYISTLITEKLSKLKLEELKTEIAKAKKCSSEFTNK 314
    |||||||
DB 80 OKIDNNNELALNNGSILAGAVYISTLITEKLSKLKLEELKTEIAKAKKCSSEFTNK 139
    |||||||
QY 315 LKSGHADLGKODATDHAKAALITKTTHATTDKAKEFKDLFESVEGILLAAOVALTNSVKE 374
    |||||||
DB 140 LKSGHADLGKODATDHAKAALITKTTHATTDKAKEFKDLFESVEGILLAAOVALTNSVKE 199
    |||||||
QY 375 LGHRNNSGGDSASTNP--DESAKGP 397
    | | | | |
DB 200 L-----TSPPVAESPCKP 212
    | | | | |

RESULT          9
US-08-158-353-4
: Sequence 4, Application US/08158353
: Patent No. 5620862
: GENERAL INFORMATION:
: APPLICANT: Padula, Steven J.
: TITLE OF INVENTION: Methods for Diagnosing Early Lyme
: TITLE OF INVENTION: Disease
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:

```

```

ADDRESSSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match 32.2%; Score 886; DB 1; Length 212;
Best Local Similarity 90.7%; Pred. No. 4e-56;
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2.

QY 195 SNSGKGDSASTPNAPDESAKGPNLTETSKRTTDSNAFVLAKEVETLVLSIDELAKKAIG 254
:|||||:
Db 20 NNSGCGDSASTPNAPDESAKGPNLTETSKRTTDSNAFVLAKEVETLVLSIDELAKKAIG 79
QY 255 OKIDNNNGIALAANNONGSILGAVAIISLTITPKLSKLNLETKETIEMAKKCSSEFTNK 314
|||||:
Db 80 OKIDNNNGIALAANNONGSILGAVAIISLTITPKLSKLNLETKETIEMAKKCSSEFTNK 139
QY 315 LKSGHADGKODATDHHAKKALIKTHATTDKCAKEFKDLFESVEGILKKAQVALTNSYKE 374
Db 140 LKSGHADGKODATDHHAKKALIKTHATTDKCAKEFKDLFESVEGILKKAQVALTNSYKE 199
QY 375 LGRNNSGDSASTNP--DESAKGP 397
|:|:|:|
Db 200 L-----TSPVVAESPCKP 212

RESULT 10
US-08-235-836C-36
Sequence 36, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-36

Query Match 32.1%; Score 882; DB 4; Length 207;
Best Local Similarity 99.5%; Pred. No. 7.4e-56;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNPDESAGNLTIVISKITDSNAFLAVKEVALLSIDELAKAIGK 438
DB 20 NNSGDSASTNPDESAGNLTIVISKITDSNAFLAVKEVALLSIDELAKAIGK 79
QY 439 NCGTLDNEANRNESILAGAYEISKLTOKLSVNSEELKKIKREAKDCSOKFTTKRSH 498
DB 80 NCGTLDNEANRNESILAGAYEISKLTOKLSVNSEELKKIKREAKDCSOKFTTKRSH 139
QY 499 AELGISVODDAAKKAIIKTHGTRKGAKELEELKSLSESKAAQALVNSVKELTNPV 558
DB 140 AELGISVODDAAKKAIIKTHGTRKGAKELEELKSLSESKAAQALVNSVKELTNPV 199
QY 559 VA 560
DB 200 VA 201

Patent 11
US-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Sautscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: burgdorferi
CURRENT APPLICATION NUMBER: US/09/196,293
FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match 31.2%; Score 858; DB 4; Length 212;
Best Local Similarity 88.8%; Pred. NO. 3.9e-54;
Matches 182; Conservative 2; Mismatches 7; Indels 14; Gaps 2;

QY 195 SNSGDSASTNPDESAGNLTIVISKITDSNAFLAVKEVALLSIDELAKAIG 254
DB 20 NNSGDSASTNPDESAGNLTIVISKITDSNAFLAVKEVALLSIDELAKAIG 79
QY 255 OKIDNNGLAALNNONGSLAGAYASTLITEKLSKNEBELKTEIAKAKCSEFTNK 314
DB 80 OKIDNNGLAALNNONGSLAGAYASTLITEKLSKNEBELKTEIAKAKCSEFTNK 139
QY 315 LKSGHADLGKODATDDHAKKAIIKTHATTDGAKAEFDLFEVGLKAAQVALTNSVKE 374
DB 140 LKSGHADLGKODATDDHAKKAIIKTHATTDGAKAEFDLFEVGLKAAQVALTNSVKE 199
QY 375 LGHRNNSGDSASTNP--DESAGP 397
DB 200 L-----TSPVVAESPGRP 212

RESULT 12
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SAUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCIT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 212
TYPE: AMINO ACID

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: US 07/903,580
3 FILING DATE: 25-JUN-1992
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: US 07/824,161
6 FILING DATE: 22-JAN-1992
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US 07/727,245
9 FILING DATE: 11-JUL-1991
10 ATTORNEY/AGENT INFORMATION:
11 NAME: BENT, Stephen A.
12 REGISTRATION NUMBER: 29,768
13 REFERENCE/DOCKET NUMBER: 30472/142 IMMU
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: (202)672-5300
16 TELEFAX: (202)672-5399
17 TELEX: 904136

```

; INFORMATION FOR SEQ ID NO: 2:
;
;   SEQUENCE CHARACTERISTICS:
;
;       LENGTH: 212 amino acids
;       TYPE: AMINO ACID
;       TOPOLOGY: linear
;
;   MOLECULE TYPE: protein
;
US-08-031-295-2

```

Query Match	24.0%:	Score 661:	DB 1:	length 212:
Best Local Similarity	73.6%:	Pred. No.	4,4e-40:	
Matches 142:	Conservative	14:	Mismatches 35:	Indels 2:
				Gaps 2:

[illegible]

RESULT 15
US-07-903-580-2
; Sequence 2, Application US/07903580
; Data No 5031353

; GENERAL INFORMATION:

APPLICANT: LIVEY, Ian

APPLICANT: DORNER, Friedrich

TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE PREVENTION OF TYPE DIARRHEA

TITLE OF INVENTION: PREVENTION OF LYME DISEASES
NUMBER OF SEQUENCES: 3

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & T

STREET: 1800 Diagonal Road

CITY: Alexandria

STATE: Virginia

; COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE

MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS

SOFTWARE: PatentIn Release #1.0
; CURRENT IDENTIFICATION CODE:

```

; CURRENT APPLICATION DATA:
APPLICATION NUMBER: FTS 007 0007 F000

```

APPLICATION NUMBER: US/01/903,580

FILED DATE: 19920623
CLASSIFICATION: 435

CEASEFICATION: 433
PRIOR APPLICATION DATA:

APPL.ICATION NUMBER:

FILING DATE: 22-JAN-1992

ISSUED DATE: 21 JAN 1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/727,245

FILING DATE: 11-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/13

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109

TELEX: 899149

```

; INFORMATION FOR SEQ ID NO: 2:
SEQUENCE OF THE CDNA:

```

```

; SEQUENCE CHARACTERISTICS:
;
;     LENGTH: 313 bp (313 bp)

```

LENGTH: 212 amino acids

1

```

;      TYPE: AMINO ACID
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-07-903-580-2

```

Query Match	24.0%	Score	661	DB	4	Length	212
Best Local Similarity	73.6%	Pred. No.	4.4e-40				
Matches	142	Conservative	14	Mismatches	35	Indels	2
						Gaps	2

OY	1	MACNNSGKDGTZ -SANSDESVKQPNLIEISKRTTDSNAVLAKAEVALLSTIDE:IAK	59
Dd	17	ISCNNSGGSGSASTINPADESARKPNLIEISKRTTDSNAFLAVAKEVTLYSSIDELATK	76
OY	60	AIGKKIHONGLDTEYHNHNSLLAGAYASTLLKÖKLDGLKNES-LKEKIDAANKCSETP	118
Dd	77	AIGKKIOONNGSLANNADNCSLLAGAYASTLITEKLALKANGSELKAKIEDAKCCSEDF	136
OY	119	TNKLEKEHTDLDGRGVDDADAKEMLKTNGTKTYGAEBLKLPFSVEVLSKAKEMLANS	178
Dd	137	TKKLACGAAOGLDIGDAVDNDNSKEALIKLTNGTKTYGAEBELVKLSBSVASLSKAQAASANS	186
OY	179	VKELTSPPVAESP 191	
Dd	197	VKELTSPPVAETP 209	

Search completed: March 18, 2002, 09:55:36
Job time: 349 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:33 ; Search time 81.02 Seconds

(without alignments)
441,544 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750
Sequence: 1 MACNNGKNGKNTSANSADSE.....KAAQALITNSVKELTPVYA 560

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

num DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCIT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep1:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	32.2	212	6	US-09-974-992-7 Sequence 7, Appl
2	878	31.9	207	6	US-09-974-992-3 Sequence 3, Appl
3	674.5	24.5	211	6	US-09-974-992-5 Sequence 5, Appl
4	234	8.5	691	5	US-09-748-875-1 Sequence 1, Appl
5	234	8.5	929	5	US-09-748-875-60 Sequence 60, Appl
6	230	8.4	707	5	US-09-748-875-2 Sequence 2, Appl
7	228	8.3	589	5	US-09-748-875-14 Sequence 14, Appl
8	223.5	8.1	690	5	US-09-748-875-61 Sequence 61, Appl
9	223.5	8.1	701	5	US-09-748-875-62 Sequence 62, Appl
10	221	8.0	711	5	US-09-748-875-3 Sequence 3, Appl
11	218.5	7.9	670	5	US-09-748-875-63 Sequence 63, Appl
12	215	7.8	1269	6	US-09-708-427-19881 Sequence 19881, A
13	212	7.7	1881	7	US-10-032-585-7646 Sequence 7646, Ap
14	212	7.7	1881	7	US-10-072-851-15590 Sequence 15590, A
15	209	7.6	1144	6	US-09-708-427-15045 Sequence 15045, A
16	209	7.6	1304	6	US-09-708-427-15045 Sequence 15045, A
17	209	7.6	1313	6	US-09-708-427-15044 Sequence 15044, A
18	201.5	7.3	1690	6	US-09-614-150-10224 Sequence 10224, A
19	201.5	7.3	1690	6	US-09-614-150-10311 Sequence 10311, A
20	197.5	7.2	805	6	US-09-708-427-3730 Sequence 3730, Ap
21	197.5	7.2	841	6	US-09-708-427-3729 Sequence 3729, Ap
22	197.5	7.2	875	6	US-09-815-242-13080 Sequence 13080, A
23	197.5	7.2	875	6	US-10-072-851-13080 Sequence 13080, A
24	197	7.2	6281	6	US-09-815-242-12996 Sequence 12996, A
25	197	7.2	6281	7	US-10-072-851-12996 Sequence 12996, A

26	193.5	7.0	2056	6	US-09-614-150-4824 Sequence 4824, Ap
27	191	6.9	2344	7	US-09-815-242-12713 Sequence 12713, A
28	191	6.9	2344	6	US-10-072-851-12713 Sequence 12713, A
29	189.5	6.9	1014	6	US-09-708-427-19883 Sequence 19883, A
30	189.5	6.9	1018	6	US-09-708-427-19882 Sequence 19882, A
31	189	6.9	730	6	US-09-708-427-3731 Sequence 3731, Ap
32	188.5	6.9	576775	4	US-08-895-611-2 Sequence 2, Appl
33	188.5	6.9	576775	4	US-08-895-611D-2 Sequence 2, Appl
34	188.5	6.9	576775	6	US-09-895-611D-2 Sequence 2, Appl
35	187	6.8	837	6	US-09-815-242-5883 Sequence 5883, Ap
36	187	6.8	1489	6	US-10-072-851-5883 Sequence 5883, Ap
37	186	6.8	1489	6	US-09-614-150-6636 Sequence 6636, Ap
38	186	6.8	1786	6	US-09-742-096-3 Sequence 3, Appl
39	185.5	6.7	3158	6	US-09-815-242-12611 Sequence 12611, A
40	185.5	6.7	3158	7	US-10-072-851-12611 Sequence 12611, A
41	183.5	6.7	2434	6	US-09-815-242-5835 Sequence 5835, Ap
42	183.5	6.7	2434	7	US-10-072-851-5835 Sequence 5835, Ap
43	183.5	6.7	5795	6	US-09-815-242-12610 Sequence 12610, A
44	183.5	6.7	5795	7	US-10-072-851-12610 Sequence 12610, A
45	182.5	6.6	1725	7	US-10-037-182-12 Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-974-992-7
; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSpC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974, 992
; PRIOR APPLICATION NUMBER: 09/180, 089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-974-992-7

Query Match 32.2%; Score 886; DB 6; Length 212;
Best Local Similarity 90.7%; Pred. No. 7.9e-49;
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;

QY 195 SNSGKGDSASTNPRADESKGRLTETISKTIIDSNFVLAWEVEVLVSTIDELAKKAIG 254
DB 20 NNSGKGDSASTNPRADESKGRLTETISKTIIDSNFVLAWEVEVLVSTIDELAKKAIG 79
QY 255 OKTIDNNNGLAALNNONGSLAGAYASTLTETKLSLKNLEELKTEIARAKKCEPFTNK 314
DB 80 OKTIDNNNGLAALNNONGSLAGAYASTLTETKLSLKNLEELKTEIARAKKCEPFTNK 139
QY 315 LKSGHADLKODATDDHAKAALIKTHATTDKGAKEFKDLEESVEGLLKAAYALTNVSK 374
DB 140 LKSGHADLKODATDDHAKAALIKTHATTDKGAKEFKDLEESVEGLLKAAYALTNVSK 199
QY 375 LGHRNNSGDSASTNP--DESAKGP 397
DB 200 L-----TSPVAESPKRP 212
RESULT 2
US-09-974-992-3

```
; Sequence 3, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSpC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-974-992-3
```

```
Query Match          31.9%; Score 878; DB 6; Length 207;
Best Local Similarity 96.9%; Pred. No. 2,4e-48;
Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNPDESANPNTLVISKTTDSNAFLAVKEVEALLSIDELSKAIGKKIK 438
      |||||||
DB 20 NNSGDSASTNPDESANPNTLVISKTTDSNAFLAVKEVEALLSIDELSKAIGKKIK 79
      |||||||

QY 439 NDGTLDNANNNESILAGAYEISKLITOKLSVNSEELKKIKKAKDCOSQKFTTKLDSDH 498
      |||||||
DB 80 NDGTLDNANNNESILAGAYEISKLITOKLSVNSEELKKIKKAKDCOSQKFTTKLDSDH 139
      |||||||

QY 499 AELGIQSVQDNNAKKAILKTGTGTRDKGAKLEELFKLSIESISKAAQALTNVSKELTNPV 558
      |||||||
DB 140 AELGIQSVQDNNAKKAILKTGTGTRDKGAKLEELFKLSIESISKAAQALTNVSKELTNPV 199
      |||||||

QY 559 VA 560
      ||
DB 200 VA 201
```

```
RESULT 3
US-09-974-992-5
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OSpC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-974-992-5
```

```
Query Match          24.5%; Score 674.5; DB 6; Length 211;
Best Local Similarity 76.7%; Pred. No. 1.6e-35;
Matches 148; Conservative 11; Mismatches 31; Indels 3; Gaps 3;
```

```
QY 1 MACNNSGKQGTNSANSADSEYKGPNTLEISKTTDSNAVLLAVEEALLSIDELTAIKA 60
      :|||||
DB 17 ISCNNSGKQGTNSANSADSEYKGPNTLEISKTTDSNAVLLAVEEALLSIDEL-TAKA 75
      :|||||

QY 61 IGGKTIHONNGDTEYNNNGSILAGAVASTLIKOKLDGLK-NEGKAKIDAAKCSFTFT 119
      |||||
DB 76 IGGKTIKNDGSLGDBANNNESILAGAYTSTLITOKLSKLSGSEGLKIAAKKCSSEFS 135
      |||||

QY 120 NKLEKHTDGLKEGVTADAKKALIKTNGT-KTGAEELKLEPSEYVLSKAAKEMLANS 178
      |||||
DB 136 TKLNDNHAQOLGQGVTDENAKKAILKANAGKDGVELEKLSGSLIESISKAAKEMLANS 195
      |||||

QY 179 VKELTSPVAESP 191
      |||||
DB 196 VKELTSPVVEESP 208
```

```
RESULT 4
US-09-748-875-1
; Sequence 1, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-1
```

```
Query Match          8.5%; Score 234; DB 5; Length 691;
Best Local Similarity 22.9%; Pred. No. 4,2e-07;
Matches 144; Conservative 94; Mismatches 260; Indels 132; Gaps 22;
```

```
QY 19 ESYK-GPNTL-----EISKTTDSNAVLLAVEEALLSIDELTAIKAIGKKIH-QNNG 70
      |||||
DB 38 EGYVSGNNLTVTSSGODISKRYAD-----EVSHELSILKQVKNLKVQHTQNVG 88
      |||||

QY 71 LDTE-----YNNNGSILAGAVASTLIKOKLDGLKNGLEKIDA 110
      |||||
DB 89 LITKLEIKKKYLYDKVAVLSEALTSKTKETKEKLAATFQFKDPLPEPEKVAEA 148
      |||||

QY 111 AKKSEFTFNKLEKHTD-----LGKE-GVTDADAKKALIKTNGTTRKGAEEIG 158
      |||||
DB 149 OKKVEAKKAKADQKDKDRNYPITTYKLELEINESDVEYKAKLELVYKAKKSQDEE 208
      |||||

QY 159 KLFES-VEVLSK-----AAKEMLANSVKELTSPVAES---PAMGSNSG 198
      |||||
DB 209 KIKQAEAEVSKQAEATRLKIKITDREKAKRADAKLEAVEKNVATSEQDKPRARAG 268
      |||||

QY 199 KCGDSASTNPADSEAKGPNTLEISKTTDSNAFVLAVEVELIVSIDELAKKAIGQ-KI 257
      |||||
DB 269 VSGELATPDKKENDAKSSD-SSVGEETLPSPDLNANNSQTEHRDVEDYIKMLSEIOL 327
      |||||

QY 258 DNNNGIAALNNONGSILAGAVASTLITEKLSKTK--NLDEL----- 297
      |||||
DB 328 DRRK-----HTQNVNLNLTKLSAIKTKYELYSVLKENSKKEELSKYAEELTAAPFQFK 382
      |||||

QY 298 -----KTELAKAKCSEEP-----TNKLSGHADLCKQDATDDHARAII 336
      |||||
DB 363 DTLKPEKRVAAEAKKVEEAKKAKADQKDEDRNYPNTTYKLELEINESDVKVAEAELEL 442
      |||||

QY 337 LKHTHTTKGAKEFDLPESEYGLKKAQVATLNSVKELGRNNSGDSASTNPDESANGK 396
      |||||
DB 443 VKEEANESRNEKIKQAKKEVES--KKAETRLKEKIT--DRKAEDEAKRAKAESEK- 496
      |||||
```

OY 397 PNLTVISKRTDSNAFLAVKEVEALLSIDELSKAIGKIKINDGTLDNENRNSLIAG 456
DB 497 --KAAEAKQKVADEAYAL-----EAKIALEVEYORLEKELEIDESSEIDYLKGLRAP 549
OY 457 AVEISKLTOKLSVNSEELKKIKKAKDCSQFTTKLDSHAELGIGSVODNAKAIL 516
DB 550 LQ--SKLDTKKRAKLSLELSKIDELDAELIAKLEVOIKDAGNNVAYFREGLEKTTA 607
OY 517 KTHGTKDKGAKLELELFKLSLESISKAQA 546
DB 608 EKKALEKAADLKKRAVDEPTAPAPQPA 637

RESULT 5
US-09-748-875-60
; Sequence 60, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 929
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-60

Query Match 8.5%; Score 234; DB 5; Length 929;
Best Local Similarity 22.9%; Pred. No. 6.3e-07;

Matches 144; Conservative 94; Mismatches 260; Indels 132; Gaps 22;

OY 19 ESWK-GPNLT-----EISKRTDSNAVLAVKEVEALLSIDELIAKAIGKIKH-QNNG 70
DB 38 EGVRSGNMLTVTSSGODISKRYAD-----EVESHLESILKIDVKKMLKKVQHTQNVG 88
OY 71 LDTE-----YHNSGLAGAYAIS-----TLKOKLDGLKNGLEKIDIDA 110
DB 89 LITKLSIEKKKYLVDLKVAVLSAEELSKTKETKEKLTATFQPKKDTLPTEPEKKVAEA 148
OY 111 AKKSEFTFNKLEKHTD-----LGKE-GVTDADAKAELIKTNGTKGAEEIG 158
DB 149 OKKVEAKKKAEDQEKEDRRNYPITYTKTLELEIASDVEYKAELELYKVAKESQDEE 208
OY 159 KLFEES-VEYLSK-----AAKEMLANSVKELTSPVAES---PAMGSNSG 198
DB 209 KIKQAEAEVESQAEATRLKIKITDREEAKKRADAKLKEAVEKNVATSEODPKRRAKRG 268
OY 199 KGGDSASTNPADSAKGNLTLSKRTDSNAFVLAVKEVEEVLVISIDELAKKAIGQ-KI 257
DB 269 VSGELATPDCKENDAKSSD-SSVGEETLPSPSLMANESQTEHRKDVEDYIKKMLSEIOL 327
OY 258 DNNNGIALNONGSLLAGAVAI STLITEKLSK---NLEEL----- 297
DB 328 DRRK-----HNGVNLNIRKLSAIKTKYELSVLKENSKEELSTKATLTAAPFOFK 382
OY 298 -----KTELAKKKCSEEF-----TNKLSGHADLGKODATDDAKAAI 336
DB 383 DTLKPEKRYAEAEKVEEKKKAKADQEKEDRRNYPITYTKTLELEIASDVEYKAELELY 442
OY 337 LKTHATTOGAEPKDLFESVEGLLKAQOVALTNSVKELGHNNSGDSASTNPDSAGK 396
DB 443 VAEENESNEEKIKQAEKVES--KKAATRLKIKIT--DRKAEAEAKKAESEK- 496
OY 397 PNLTVISKRTDSNAFLAVKEVEALLSIDELSKAIGKIKINDGTLDNENRNSLIAG 456

DB 497 --KAAEAKQKVADEAYAL-----EAKIALEVEYORLEKELEIDESSEIDYLKGLRAP 549
OY 457 AVEISKLTOKLSVNSEELKKIKKAKDCSQFTTKLDSHAELGIGSVODNAKAIL 516
DB 550 LQ--SKLDTKKRAKLSLELSKIDELDAELIAKLEVOIKDAGNNVAYFREGLEKTTA 607
OY 517 KTHGTKDKGAKLELELFKLSLESISKAQA 546
DB 608 EKKALEKAADLKKRAVDEPTAPAPQPA 637

RESULT 6
US-09-748-875-2
; Sequence 2, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-2

Query Match 8.4%; Score 230; DB 5; Length 707;
Best Local Similarity 22.2%; Pred. No. 7.7e-07;

Matches 139; Conservative 100; Mismatches 264; Indels 122; Gaps 21;

OY 19 ESWK-GPNLT-----EISKRTDSNAVLAVKEVEALLSIDELIAKAIGKIKH-QNNG 70
DB 38 EGVRSGNMLTVTSSGODISKRYAD-----EVESHLESILKIDVKKMLKKVQHTQNVG 88
OY 71 LDTE-----YHNSGLAGAYAIS-----TLKOKLDGLKNGLEKIDIDA 110
DB 89 LITKLSIEKKKYLVDLKVAVLSAEELSKTKETKEKLTATFQPKKDTLPTEPEKKVAEA 148
OY 111 AKKSEFTFNKLEKHTD-----LGKE-GVTDADAKAELIKTNGTKGAEEIG 158
DB 149 OKKVEAKKKAEDQEKEDRRNYPITYTKTLELEIASDVEYKAELELYKVAKESQDEE 208
OY 159 KLFEES-VEYLSK-----AAKEMLANSVKELTSPVAES---PAMGSNSG 198
DB 209 KIKQAEAEVESQAEATRLKIKITDREEAKKRADAKLKEAVEKNVATSEODPKRRAKRG 268
OY 199 KGGDSASTNPADSAKGNLTLSKRTDSNAFVLAVKEVEEVLVISIDELAKKAIGQ--- 255
DB 269 VSGELATPDCKENDAKSSD-SSVGEETLPSPSLMANESQTEHRKDVEDYIKKMLSEIOL 327
OY 256 -----KIDNNGIALA-----NNONGSLAGAYAI STLITEKLSK----- 290
DB 328 DRRKTPNVLNIRKLSAIKTKYELSVLKENSKEELSTKATLTAAPFOFKDTLP 387
OY 291 LKMLELTKELIAKKKCSSEF-----TNKLSGHADLGKODATDDAKAAI LKTHA 341
DB 388 EKKVAEAEKVEEKKKAKADQEKEDRRNYPITYTKTLELEIASDVEYKAELELYKEEA 447
OY 342 TTDGAKPEKDLFESVEGLLKAQOVALTNSVKELGHNNSGDSASTNPDSAKGNLT 401
DB 448 NESRNEEKIKQAEKVES--KKAATRLKIKIT--DRKAEAEAKKAESEK---KAA 499
OY 402 ISKRTDSNAFLAVKEVEALLSIDELSKAIGKIKINDGTLDNENRNSLIAGAYEIS 461
DB 500 EAKQKVADEAYAL-----EAKIALEVEYORLEKELEIDESSEIDYLGGLAPLQ--S 552
OY 462 KLTQKLSVNSEELKKIKKAKDCSQFTTKLDSHAELGIGSVODNAKKAILTHTGT 521

Db 553 KLTGKAKKLSLELSKIDELDAEIAKLEVLQKDAEAGNNVAVYFKEGLEKTTAKKAE 612
QY 522 KDKGAKLELEFKLSLSKAAQAA 546
Db 613 LEKAEADLKRAVDEPETPAPAPQA 637

RESULT 7

US-09-748-875-14
; Sequence 14, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 589
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-14

Query Match 8.3%; Score 228; DB 5; Length 589;
Best Local Similarity 23.6%; Pred. No. 8.1e-07;
Matches 148; Conservative 89; Mismatches 247; Indels 142; Gaps 24;

QY 19 ESVK-GPNLT-----EISKRTIDSNVLAFLAVEEALLSIDETIAKAIGKIH-QNNG 70
Db 1 EGVRSGNMLVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 51
QY 71 LDTE-----YHNSSLAGAYAIS-----TLIKOKLDLKNEGLEKIDA 110
Db 52 LITKLSLKKKYLVDLKVNLSEALTSKTEKELTAFEFQKDTLPTEPEKVAEA 111
QY 111 AKKSEFTNKLKLEKHTD-----LGKE-GVTDADAKKALIKTNGTKGAELG 158
Db 112 OKKVEAKKKAEDOKERRRNPITYTKTLELTAESDVEVKKAELELVKAKESQDEE 171
QY 159 KLFEES-VEVLSK-----AAKEMLANGVELTSPVAVES-----PAMGSNG 198
Db 172 KIKQAEAEVESKQAEATRLKIKITDREAKRKADAKLEAVEKNVATSEODKPRRAKRG 231
QY 199 KGGDSASTNPADSAKGNLTETSKITDSNAFLAVKEVETLVLSIDELAKKAIGQ-KI 257
Db 232 VSGELATPPDKKENDAKSSD--SSVGEETLPPSLMANESQTEHRKDVDEYIKKMLSEIQL 290
QY 258 DNNNGLAALNONGSLAGAYASTILTEKLSKLK-----NLEEL----- 297
Db 291 DRRK-----HTQVNNLNINIKLSAITKYLYELSVLKENSKEELTSKTAELTAAEQFKK 345
QY 298 -----KTEIAKAKKCEEF-----TNKLKSGHADLKGODATDHAKAAL 336
Db 346 DTLKPEKVAEAEKKVEAKKKAADOKEDRRNPITYTKTLELTAESDVEVKKAELELV 405
QY 337 LKTHATDCAKAEKFDLFESVEGLKAAQVALNSVKEIGHRNNSGDSASTNPDESAGK 396
Db 406 VKEAANSREBEKIKAKKEVES--KKAETRLKIKIT--DRKKAEEKARRAESSEK- 459
QY 397 PNLTIVISKRTIDSNAPFLAVKEVEALLSIDELSKAIGKIKRNDGTLDNANRNESLNG 456
Db 460 --KAAEAKQVDAEEVAL-----EAKIALEIYEVQLEKLEKIDESDSEDIYKBEGLRAP 512
QY 457 AYVISTLIQKLSVNSSEELKKIKAKDCSQFTTKLDSHAELIGQSVODNNAKAIL 516
Db 513 LQ--SKLDTKRAKLSKLELSKIDELDAEIAKLEVLQDA-----EGNNNVEAVEF 561

QY 517 KTHGTOKGAKLELEFKLSLSLSKA 542
Db 562 K-EGLEKTTAEKKALEKAEADLKRA 586

RESULT 8

US-09-748-875-61
; Sequence 61, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-61

Query Match 8.1%; Score 223.5; DB 5; Length 690;
Best Local Similarity 22.7%; Pred. No. 1.9e-06;
Matches 143; Conservative 92; Mismatches 263; Indels 131; Gaps 22;

QY 19 ESVK-GPNLT-----EISKRTIDSNVLAFLAVEEALLSIDETIAKAIGKIH-QNNG 70
Db 38 EGVRSGNMLVTSSGODISKRYAD-----EVESHLESILKDVKNLKKVQHTQNVG 88
QY 71 LDTE-----YHNSSLAGAYAIS-----TLIKOKLDLKNEGLEKIDA 110
Db 89 LITKLSLKKKYLVDLKVNLSEALTSKTEKELTAFEFQKDTLPTEPEKVAEA 148
QY 111 AKKSEFTNKLKLEKHTD-----LGKE-GVTDADAKKALIKTNGTKGAELG 158
Db 149 OKKVEAKKKAEDOKERRRNPITYTKTLELTAESDVEVKKAELELVKAKESQDEE 208
QY 159 KLFEES-VEVLSK-----AAKEMLANGVELTSPVAVES-----PAMGSNG 198
Db 209 KIKQAEAEVESKQAEATRLKIKITDREAKRKADAKLEAVEKNVATSEODKPRRAKRG 268
QY 199 KGGDSASTNPADSAKGNLTETSKITDSNAFLAVKEVETLVLSIDELAKKAIGQ-KI 257
Db 269 VSGELATPPDKKENDAKSSD--SSVGEETLPPSLMANESQTEHRKDVDEYIKKMLSEIQL 327
QY 258 DNNNGLAALNONGSLAGAYASTILTEKLSKLK-----NLEEL----- 297
Db 328 DRRK-----HTQVNNLNINIKLSAITKYLYELSVLKENSKEELTSKTAELTAAEQFKK 382
QY 298 -----KTEIAKAKKCEEFNTNKLKSGHAD-----LGKODATDHAKAAL 337
Db 383 DTLKPEKVAEAEKKVEAKKKAADOKEDRRNPITYTKTLELTAESDVEVKKAELELV 442
QY 338 KTHATDCAKAEKFDLFESVEGLKAAQVALNSVKEIGHRNNSGDSASTNPDESAGK 397
Db 443 KEAANSREBEKIKAKKEVES--KKAETRLKIKIT--DRKKAEEKARRAESSEK- 495
QY 398 NLTIVISKRTIDSNAPFLAVKEVEALLSIDELSKAIGKIKRNDGTLDNANRNESLNG 457
Db 496 -KAAEAKQVDAEEVAL-----EAKIALEIYEVQLEKLEKIDESDSEDIYKBEGLRAP 549
QY 458 YVISTLIQKLSVNSSEELKKIKAKDCSQFTTKLDSHAELIGQSVODNNAKAIL 517
Db 550 Q--SKLDTKRAKLSKLELSKIDELDAEIAKLEVLQDAEAGNNNVEAVEYFKEGLEKTTAE 607
QY 518 THTGTDKGAKELELEFKLSLSKAAQAA 546
Db 608 KKALEKAEADLKRAVDEPETPAPAPQA 636


```

; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-748-875-63

```

Query Match 7.9%; Score 218.5; DB 5; Length 670;
 Best Local Similarity 22.8%; Pred. No. 3.8e-06;
 Matches 142; Conservative 81; Mismatches 248; Indels 151; Gaps 23;

```

QY 12 TSANSADESVKGNPLTEISKITDSNAVLAVKEVEALLSIDELIAAKAIG-----KTIH 66
DB 48 TSYKNANES-----OTENRK-----AAQVDEDIKMLSEIOEYIKKMLSEIODKRRH 96
QY 67 QNNGLDTEYNHNGSLIAGYAISTLIKOKLDGKNEGLEKIDA-----AKKCSB-----T 117
DB 97 TON-----VNLNRKL-----SAIQTKLYELRYLKEKSKKEELTSVYPTNTYKTLLELEIAE 147
QY 118 FNKLKEKHKTGK-----EGVTDADAKKAILKTGCT-----KTYGAELGKLF 161
DB 148 FVYKAKKEALELVEKAEAPRNEKIKOKAKVESKKAATRLTEIKTERKKAEEAKKKA 207
QY 162 ESVEVLISRAAKEMLANSVK-----ELTSPVVAESPAMSGSGGD---SAS 205
DB 208 BESEKKAALAEKQVDTKQKPKRRAKRGVSGELATPKKENDAKSSDSVGEELTSPS 267
QY 206 TNPAPESAKGNPLTEISKITDSNAFVLAKEVEETLVISIDELAKKAIGQ-KIDNNNGLA 264
DB 268 LMANESO-----TEHRR-----DVEYIKKMLSEIOLDRRK--- 299
QY 265 ALNNGSLAGAYAISTLITEKLSKIK---NLEEL-----KT 299
DB 300 ---HQNVNLNIRKLSAIKTKYLYELSVLEKNSKREBELSKTAELTAPEOKKDTLPK 357
QY 300 EIAAKKCSSEFTTKLKSAD-----LGKODATDDHAKAAILKTHATTD 344
DB 358 KVAARAKKVEEAKKRAKDKQEDERNTPTNTYKTLLEAESDVVDAKAELELVKEANES 417
QY 345 KGAKEFDLFEVSGELKKAQOVALTNSVKELGHNNNGSGSASTNPDESAGPNLVISK 404
DB 418 RNEKIKQAKKEKVES--KKAETATLEKIKT--DDRKAEEEAARRKAESEK---KAAEAK 470
QY 405 KIYDSNAFLAVKEVEALLSIDELSKAIGKIKINDGTLDNEANRNESILAGAVEISKLI 464
DB 471 OKVDAEYAL-----EAKIAELEVEYQRLKEKLEIDESSEDIYKGLAPAO---SKLD 523
QY 465 TOKLSVNSESEIKKIKKEANDCSOKFTTKLDSIAELGISOVDONAKKAILKTHGTGDK 524
DB 524 TKRAKLSKLELSKIDELDAETAKLEVOULKDAEGNNNVAEYFKEGLEKTAJAEKAELEK 583
QY 525 GAKLELELEFSLKSLKAAQAA 546
DB 584 AEADLKAAVDEPAPAPQPA 605

```

RESULT 12
 US-09-708-427-19881

; Sequence 19881, Application US/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; TITLE OF INVENTION: THEREBY

```

; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..1269
; OTHER INFORMATION:
; US-09-708-427-19881

```

Query Match 7.8%; Score 215; DB 6; Length 1269;
 Best Local Similarity 22.1%; Pred. No. 1.5e-05;
 Matches 155; Conservative 118; Mismatches 255; Indels 172; Gaps 33;

```

QY 13 SANGAD-ESYKGNPLTEISKITDSNAVLAVKEVEALLSIDELIAAKAIGKTIHONNGL 71
DB 224 SASADSESKALEFSELSTKES-----AKEMEKMAISLOO-EIKLEKNEKSENEKV 276
QY 72 DTEYNHNGSLIAGYAISTLIKOKLDGKNEGLEKIDAKKCSSEFTTKLKEKHKTG 130
DB 277 EALAKSSAGELAAVOEELALSKSRL--LETE---QWVSTFALIDELTQLEBKRRKASES 331
QY 131 -KEGVY---DADAKKAILKT-----NCTKTKGABELKLFESVEVLISRAAKEMLANSVK 181
DB 332 FKEELSVLODLDQVTKGLQAKLSEQEGINSKLAEL--KEKELLESLSKQDEKTL--RTANE 389
QY 182 LTPVVAESPAMSGSGDSAST-----NPDESAGK-----PILNE 220
DB 390 KLAEVLKEREKALBANVAEYTSNVAITVENCNELEKELKTSDENFSTDALISQALSNNSE 449
QY 221 ISKTIID-----SNAYLAVK---EVEYTLVISIDELAKKAIGKIDNNNGLA 269
DB 450 LEKLSLELSEHSEAGSAAATAATOKNMLELDVYRSSQAEEAKSQIKLETFKTAABEQ 509
QY 270 NCSILAGAYAISTLITEKLSKIKNEBELKTEIATAKAKKCSSEFTN-----KLKSGHA 320
DB 510 NAELEQOINILQKSDAREBELKESEKSELOTALEVAEEKKOATTOQMOEYKQKASEL 569
QY 321 DL-----GKQDATDHAKAAILK--THA---TTDKGAKKFDL-----FESVEGL 361
DB 570 ELSTIOSARNSELEEDLRIALOKGAHEHEDRANTTHORSIELEGLQSSQSHDEAEGRL 629
QY 362 KAAOVALTNS---VKELGRNNS-----GGDS----- 385
DB 630 KOLELLQTEKRIQLEOVSVSLKKEHGETBADSGLYGOVAELQSLYLEAFVYKSSSLE 689
QY 386 -----ASTNPDESAGPNLVISKITDS--NAFLAVKEVEALLSI--DELSKAIK-- 435
DB 690 AALNITATENEKELTENLNAVTESEKKLEATVDEYVKSISESNLLESTIRNELNTYQGLE 749
QY 436 KIKND---GTIDNPA---NRNLSLAGAVEISKLI-----OKLSVNSEELK 477
DB 750 STENDLKAAGLOEISEVMERKLSAESELEOKGREIDEATYTKMELEALHQSISIDSEHRLQ 809
QY 478 KKIK-----EAKDCOKFTT---KLKDSHAELGISOVDONAKKAILKTHG----- 520
DB 810 KAMEEFTSRDSEASSLFTKLDLBCKIKSYFEQOLAESAAGKSSLSKEKLEOTLGLRLAAES 869
QY 521 TKDKAKLELEL--TKSLSLKAAQAAITNS-----VKEL 554
DB 870 VNEKLEQFDOAKESLOSSESSEILAEITNNQLKIKIOEL 909

```

RESULT 13

```

US-10-032-585-7646
: Sequence 7646, Application US/10032585
: GENERAL INFORMATION:
: APPLICANT: Terry, Roemer D.
: APPLICANT: Bo, Jlang
: APPLICANT: Charles, Boone
: APPLICANT: Howard, Bussey
: TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
: FILE REFERENCE: 10182-005-999
: CURRENT APPLICATION NUMBER: US/10/032,585
: CURRENT FILING DATE: 2001-12-20
: NUMBER OF SEQ ID NOS: 8000
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 7646
: LENGTH: 1881
: TYPE: PRT
: ORGANISM: Candida albicans
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1881)..(1881)
: OTHER INFORMATION: X-any amino acid
: 10-032-585-7646

```

Db	492	SL-----VDEATKLIQSIQSEBIDLKREKAGVLYKIIIEELSVANESLADNYDDLOSIYESK	544
Qy	361	LKAQVAVLITNSVLEIGHRRNSGDSAS--TNPESA-----KGNITVYSKK	405
Db	547	DLKEEVAYLYKIIIEELSVANESLVDRKTKLOHIDQEAEBELGREAHLKIIIEELSEKENEN	606
Qy	406	ITDSNAFLLAWE-----ALLSIDELSKAIKKIKINQDTLNEANRNE-----	451
Db	607	LYDVNVAQNQIAESKDLREREVAYLYKIIDELSTANSTLADNYNTINQISENKELERE	666
Qy	452	SLIAQVEISLITOKLIVSNPELKKRIKKAQDSOKFTTKIXDHAELQISQVDDN	510
Db	667	TLTLKKAEBSEL--NSIYDKAKLQTLVQENELRERETAYLYKIIIEELSKIIIEELISDO	724
Qy	511	AKKALITTHGKCKGADELFL--KSLIESLSKAQAALITNSVKELTNPV	559
Db	725	ETKLIQISNHEKEE--LKERETAAYLYKIIIEELSK--VQEDLLKREKELTGMVV	771

Db 725 ETKLQISNHEKEE--LKERETAYLKKEI EELSK-VQEDDLNKENELHGMV 771

Db 725 ETKQISNHEKEE-LKRETAIYIKITEELSK-VQEDLNKENELHGMVV 771
Search completed: March 18, 2002, 09:58:37
Job time: 395 sec

., Tue Mar 19 10:57:58 2002

us-09-596-746a-52.rapn

Page 9

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:49 ; Search time 68.77 Seconds

(without alignments)
620.296 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750

Sequence: 1 MACNNSGKQGNSTANSANADES.....KAAQAALITNSKELTNPVA 560

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	33.7	210	2	G70218
2	921	33.5	210	2	S69927
3	886	32.2	212	2	S20543
4	881	32.0	207	2	S37727
5	878	31.9	207	2	S69919
6	877	31.9	212	2	S69921
7	874	31.8	200	2	I40122
8	853	31.0	177	2	I40129
9	846	30.8	179	2	S54187
10	844	30.7	179	2	S54188
11	834.5	30.3	180	2	S54191
12	832	30.3	179	2	S54186
13	811	29.5	189	2	S70263
14	798	29.0	207	2	I40271
15	794	28.9	200	2	S54198
16	786	28.6	207	2	I40276
17	759.5	27.6	199	2	S54197
18	757.5	27.5	211	2	I40145
19	756.5	27.5	193	2	S70279
20	700	25.5	210	2	I40144
21	696.5	25.3	193	2	S70287
22	683.5	24.9	211	2	I40277
23	677.5	24.6	201	2	S69918
24	676.5	24.6	209	2	I40273
25	674.5	24.5	191	2	S70278
26	674.5	24.5	191	2	I40153
27	673.5	24.5	191	2	I40268
28	672.5	24.5	193	2	S70280
29	672	24.4	212	2	S69922

30	667	24.3	194	2	S70268	outer surface prot
31	666.5	24.2	193	2	S70265	outer surface prot
32	666.5	24.2	211	2	I40278	outer surface prot
33	666	24.2	194	2	S70277	outer surface prot
34	665.5	24.2	209	2	S69917	outer surface prot
35	665	24.2	190	2	S70273	outer surface prot
36	661	24.0	212	2	S70254	outer surface prot
37	657.5	23.9	193	2	S70276	outer surface prot
38	656.5	23.9	193	2	S70274	outer surface prot
39	653.5	23.8	180	2	S54189	outer surface prot
40	653.5	23.8	191	2	S70284	outer surface prot
41	652.5	23.7	211	2	S69932	outer surface prot
42	646	23.5	212	2	I40279	outer surface prot
43	642.5	23.4	209	2	I40142	outer surface prot
44	635	23.1	212	2	I40143	outer surface prot
45	634.5	23.1	203	2	I40108	outer surface prot

ALIGNMENTS

RESULT: 1
G70218
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C:Accession: G70218; I40269; S37726; S70281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.C.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: G70218
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <KLE>
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66339.1; PID:g2689901; TIGR:BB
A:Experimental source: strain B31
R:Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: I40269; MUID:96025162
A:Accession: I40269
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:DA9497; NID:g707092; PIDN:BA008457.1; PID:g765684
R:Jauris-Helpe, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (Os
A:Reference number: S37726; MUID:93268136
A:Accession: S37726
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <JAD>
A:Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R:Rilvey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; MUID:96296448
A:Accession: S70281
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210 <LIV>
A:Cross-references: EMBL:I42887; NID:g858715; PIDN:AB36995.1; PID:g1695212
A:Experimental source: strain Ip2
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 33.7%; Score 928; DB 2; Length 210;
 Best Local Similarity 98.4%; Pred. No. 1.6e-34;
 Matches 186; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTSANSADSEYKGNLTLEISKTTDSNAVLAVKEVALLSIDELAIKA 60
 Db 17 ISCNNSKDGKTSANSADSEYKGNLTLEISKTTDSNAVLAVKEVALLSIDELAIKA 76
 QY 61 IGGKTIHONNGLDTEYNNHNGSLAGAVAI STLTKOKLDGLKNEGLKEKIDAKKCEFTFN 120
 Db 77 IGGKTIHONNGLDTEYNNHNGSLAGAVAI STLTKOKLDGLKNEGLKEKIDAKKCEFTFN 136
 QY 121 KLRKHTDLCKEGVTDADAKAEALIKTNGTKGAEEIGKLFESVEVLSKRAKEMLANSVK 180
 Db 137 KLRKHTDLCKEGVTDADAKAEALIKTNGTKGAEEIGKLFESVEVLSKRAKEMLANSVK 196
 QY 181 ELTSPVVAESP 191
 Db 197 ELTSPVVAESP 207

RESULT 2

outer surface protein C precursor - Lyme disease spirochete (strain PKA)
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 A:Variate: strain PKA
 C:Date: 15-Feb-1997 #sequence-revision 27-Feb-1997 #text-change 26-May-2000
 C:Accession: S69927; S72669
 R:Jauris-Heipke, S.; Lieg1, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek, E.
 J. Clin. Microbiol. 33, 1860-1866, 1995
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
 A:Reference number: 140047; MUID:95395018
 A:Accession: S69927
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-210 <JAU>
 A:Cross-references: EMBL:X69589
 A:Experimental source: strain PKA
 R:Jauris, S.
 submitted to the EMBL Data Library, February 1994
 A:Reference number: S72669
 A:Accession: S72669
 A:Molecule type: DNA
 A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAM>
 A:Cross-references: EMBL:X69589
 C:Genetics:
 A:Gene: ospC
 A:superfamily: Lyme disease spirochete surface protein C

Query Match 33.5%; Score 921; DB 2; Length 210;
 Best Local Similarity 97.4%; Pred. No. 3.2e-34;
 Matches 186; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGTSANSADSEYKGNLTLEISKTTDSNAVLAVKEVALLSIDELAIKA 60
 Db 17 ISCNNSKDGKTSANSADSEYKGNLTLEISKTTDSNAVLAVKEVALLSIDELAIKA 76
 QY 61 IGGKTIHONNGLDTEYNNHNGSLAGAVAI STLTKOKLDGLKNEGLKEKIDAKKCEFTFN 120
 Db 77 IGGKTIHONNGLDTEYNNHNGSLAGAVAI STLTKOKLDGLKNEGLKEKIDAKKCEFTFN 136
 QY 121 KLRKHTDLCKEGVTDADAKAEALIKTNGTKGAEEIGKLFESVEVLSKRAKEMLANSVK 180
 Db 137 KLRKHTDLCKEGVTDADAKAEALIKTNGTKGAEEIGKLFESVEVLSKRAKEMLANSVK 196
 QY 181 ELTSPVVAESP 191
 Db 197 ELTSPVVAESP 207

RESULT 3
 S20543

outer surface protein C precursor - Lyme disease spirochete
 M:Alternate names: membrane-associated protein PC
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Jan-1995 #sequence-revision 13-Jan-1995 #text-change 21-Jul-2000
 C:Accession: S20543; I40107; S69931; S70275; S34175
 R:Fuchs, R.; Jauris, S.; Lottspeich, F.; Preac-Mursic, V.; Wilske, B.; Soutschek, E.
 Mol. Microbiol. 6, 503-509, 1992
 A:Title: Molecular analysis and expression of a Borrelia burgdorferi gene encoding a
 A:Reference number: S20543; MUID:92219995
 A:Accession: S20543
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-212 <FUC>
 A:Cross-references: EMBL:X62162; NID:933932; PIDN:CAA44093.1; PID:933933
 R:Thiesen, M.; Frederiksen, B.; Lebeck, A.M.; Vuust, J.; Hansen, K.
 J. Clin. Microbiol. 31, 2570-2576, 1993
 A:Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os
 A:Reference number: 140105; MUID:94075528
 A:Accession: I40107
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-205 <RES>
 A:Cross-references: EMBL:X73624; NID:9313273; PIDN:CAA52003.1; PID:9313274
 A:Experimental source: strain DK26
 R:Jauris-Heipke, S.; Lieg1, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek
 J. Clin. Microbiol. 33, 1860-1866, 1995
 A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
 A:Reference number: 140047; MUID:95395018
 A:Accession: S69931
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-212 <JAU>
 A:Cross-references: EMBL:X69590
 A:Experimental source: strain Pwudi
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1994
 R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.
 Mol. Microbiol. 18, 257-269, 1995
 A:Title: Evidence for lateral transfer and recombination in ospC variation in Lyme di
 A:Reference number: S70255; MUID:96296448
 A:Accession: S70275
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 19-212 <LIV>
 A:Cross-references: EMBL:L42883; NID:9858709; PIDN:AAB36991.1; PID:91695209
 A:Experimental source: strain JSB
 C:Genetics:
 A:Gene: ospC
 A:superfamily: Lyme disease spirochete surface protein C

Query Match 32.2%; Score 886; DB 2; Length 212;
 Best Local Similarity 90.7%; Pred. No. 1.1e-32;
 Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;

QY 195 SMSGKGDASSTNPDESAGKGNLTLEISKTTDSNAVLAVKEVALLSIDELAIKA 254
 Db 20 NNSGKGDASSTNPDESAGKGNLTLEISKTTDSNAVLAVKEVALLSIDELAIKA 79
 QY 255 OKIDNNNGLAALNNGNSLGAAYASTLTKOKLDGLKNEGLKEKIDAKKCEFTFN 314
 Db 80 OKIDNNNGLAALNNGNSLGAAYASTLTKOKLDGLKNEGLKEKIDAKKCEFTFN 139
 QY 315 LKSGHADLGKODATDDHAKAALIKTHATTDGAKFEKDLFESVEGLKAAQVALTNSVKE 374
 Db 140 LKSGHADLGKODATDDHAKAALIKTHATTDGAKFEKDLFESVEGLKAAQVALTNSVKE 199
 QY 375 LGHRNNSGDSASTNP--DESAKGP 397
 Db 200 L-----TSPVVAESP 212

RESULT 4
 S37727

outer surface protein C precursor - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-May-2000

C:Accession: S37727

R:Jauris-Helpeke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W

Med. Microbiol. Immunol. 182, 37-50, 1993

A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)

A:Reference number: S37726; MUID:93268136

A:Accession: S37727

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-207 <JAU>

A:Cross-references: EMBL:X69595; NID:g311393; PIDN:CAA5305.1; PID:g311394

C:Superfamily: Lyme disease spirochete surface protein C

Query Match

Best Local Similarity 32.0%; Score 881; DB 2; Length 207;

Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNDESAKGNPLTVISKKTIDSNAPFLAVKEVALLSSIDELSKAIGKIK 438

DB 20 NNSGDSASTNDESAKGNPLTVISKKTIDSNAPFLAVKEVALLSSIDELSKAIGKIK 79

QY 439 NNGTLDNEANRRESLIAGAYEISKLTOKLSVLNSELKKEKKEADCSOKFTTKLDKSH 498

DB 80 NNGTLDNEANRRESLIAGAYEISKLTOKLSVLNSELKKEKKEADCSOKFTTKLDKSH 139

QY 499 AELGIQSVODDNNAKKAILTKHTGKDGAKELFELFKSLSKAAQALTNVSKELTNV 558

DB 140 AELGIQSVODDNNAKKAILTKHTGKDGAKELFELFKSLSKAAQALTNVSKELTNV 199

QY 559 VA 560

DB 200 VA 201

RESULT 5

S69919

outer surface protein C precursor - Borrelia garinii (strain PTrob)

C:Species: Borrelia garinii

A:Variety: strain PTrob

C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000

C:Accession: S69919

R:Jauris-Helpeke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E

J. Clin. Microbiol. 33, 1860-1866, 1995

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia

A:Reference number: 140047; MUID:95395018

A:Accession: S69919

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-207 <JAU>

A:Cross-references: EMBL:X83554; NID:g872027; PIDN:CAA58544.1; PID:g872028

A:Experimental source: strain PTrob

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

F.1-18/Domain: signal sequence #status predicted <SIG>

F.19-207/Product: outer surface protein C #status predicted <MAT>

Query Match

Best Local Similarity 31.9%; Score 878; DB 2; Length 207;

Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NNSGDSASTNDESAKGNPLTVISKKTIDSNAPFLAVKEVALLSSIDELSKAIGKIK 438

DB 20 NNSGDSASTNDESAKGNPLTVISKKTIDSNAPFLAVKEVALLSSIDELSKAIGKIK 79

QY 439 NNGTLDNEANRRESLIAGAYEISKLTOKLSVLNSELKKEKKEADCSOKFTTKLDKSH 498

DB 80 NNGTLDNEANRRESLIAGAYEISKLTOKLSVLNSELKKEKKEADCSOKFTTKLDKSH 139

QY 559 VA 560

DB 200 VA 201

RESULT 6

S69921

outer surface protein C precursor - Borrelia afzelii (strain PBO)

C:Species: Borrelia afzelii

A:Variety: strain PBO

C:Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000

C:Accession: S69921; 140047

R:Jauris-Helpeke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek

J. Clin. Microbiol. 33, 1860-1866, 1995

A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre

A:Reference number: 140047; MUID:95395018

A:Accession: S69921

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-212 <JAU>

A:Cross-references: EMBL:X81521; NID:g804954; PIDN:CAA57241.1; PID:g804955

A:Experimental source: strain PBO

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

F.1-18/Domain: signal sequence #status predicted <SIG>

F.19-207/Product: outer surface protein C #status predicted <MAT>

Query Match

Best Local Similarity 31.9%; Score 877; DB 2; Length 212;

Matches 184; Conservative 4; Mismatches 3; Indels 14; Gaps 2;

QY 195 NNSGDSASTNDESAKGNPLTVISKKTIDSNAPFLAVKEVALLSSIDELSKAIG 254

DB 20 NNSGDSASTNDESAKGNPLTVISKKTIDSNAPFLAVKEVALLSSIDELSKAIG 79

QY 255 OKIDNNNGLAALNONGSLAGAYASTLTETKSLKLNLEELKTEIAKAKCSEFTNK 314

DB 80 OKIDNNNGLAALNONGSLAGAYASTLTETKSLKLNLEELKTEIAKAKCSEFTNK 139

QY 315 LKSGHADLCKODATDHAHAAILTKHTGKDGAKELFELFKSLSKAAQALTNVSK 374

DB 140 LKSGHADLCKODATDHAHAAILTKHTGKDGAKELFELFKSLSKAAQALTNVSK 199

QY 375 LGHRNNSGDSASTNDESAKGP 397

DB 200 LGHRNNSGDSASTNDESAKGP 212

RESULT 7

140122

outer surface protein C - Lyme disease spirochete (strain DK6) (fragment)

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000

C:Accession: I40122; S34177

R:Thiesen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.

J. Clin. Microbiol. 31, 2570-2576, 1993

A:Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os

A:Reference number: I40105; MUID:94075528

A:Accession: I40122

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-200 <RES>

A:Cross-references: EMBL:X73626; NID:g313277; PIDN:CAA52005.1; PID:g313278

C:Genetics:

A:Gene: ospC

C:Superfamily: Lyme disease spirochete surface protein C

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:57 ; Search time 39.62 Seconds

(without alignments)
518.231 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750
Sequence: 1 MACNNGKNGKNTSANSANDES.....KAAQAALTNVKELTNPVVA 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	928	33.7	210	1	OSCL_BORBU
2	886	32.2	212	1	OSCL_BORBU
3	402.5	14.6	214	1	VM24_BORHE
4	398	14.5	215	1	VM03_BORHE
5	210.5	7.7	1957	1	Y086_SCHPO
6	207.5	7.5	1790	1	US01_YEAST
7	207.5	7.5	2869	1	RBP1_PLAVB
8	204.5	7.4	1509	1	MYSN_ACACA
9	193.5	7.0	2017	1	MYSN_DROME
10	193.5	7.0	2116	1	MYS2_DICDI
11	193	7.0	1433	1	REST_CHICK
12	192	7.0	1966	1	MYSB_CAEL
13	189	6.9	1022	1	SCA4_RICPR
14	189	6.9	1427	1	REST_HUMAN
15	187	6.8	875	1	ZIP1_YEAST
16	187	6.8	1164	1	BAG_STRAG
17	186	6.8	1167	1	CAGA_HELPJ
18	185	6.7	1189	1	SCIL_CHICK
19	185	6.7	1251	1	RBP2_PLAVB
20	184.5	6.7	1636	1	BUD3_YEAST
21	183.5	6.7	1189	1	YJH6_YEAST
22	183.5	6.7	1940	1	MYH3_CHICK
23	182.5	6.6	1786	1	LMB1_MOUSE
24	180.5	6.6	775	1	YHGE_BACSU
25	180	6.5	3672	1	LML2_CAEL
26	179.5	6.5	2411	1	MYS4_DROME
27	179	6.5	1940	1	MYH3_RAT
28	178.5	6.5	727	1	MPPL_ARATH
29	178.5	6.5	1290	1	XCPQ_XENLA
30	177.5	6.5	1786	1	LMB1_HUMAN
31	177	6.4	1875	1	MLP1_YEAST
32	177	6.4	2748	1	NUM1_YEAST
33	176.5	6.4	998	1	SCA4_RICAK

ALIGNMENTS

ALIGNMENTS				
RESULT	ID	OSCL_BORBU	STANDARD:	PRT: 210 AA.
1	OSCL_BORBU	00737;	15-DEC-1998 (Rel. 37, Created)	
AC	DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=93268136; Pubmed=8098841;		
RX	RA	Jauris-Helpe S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,		
RA	RA	Will G., Wilske B.;		
RT	RT	"Genetic heterogeneity of the genes coding for the outer surface		
RT	RT	protein C (OspC) and the flagellin of Borrelia burgdorferi.";		
RL	RL	Med. Microbiol. Immunol. 182:37-50(1993).		
RN	[2]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=93239332; Pubmed=8478108;		
RX	RA	Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,		
RA	RA	Schwab E., Wanner G.;		
RT	RT	"Immunological and molecular polymorphisms of OspC, an immunodominant		
RT	RT	major outer surface protein of Borrelia burgdorferi.";		
RL	RL	Infect. Immun. 61:2182-2191(1993).		
RN	[3]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96025162; Pubmed=7494039;		
RX	RA	Fukunaga M., Hamase A.;		
RA	RA	"Outer surface protein C gene sequence analysis of Borrelia		
RT	RT	burgdorferi sensu lato isolates from Japan.";		
RL	RL	Infect. Immun. 61:5097-5105(1993).		
RN	[4]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96025162; Pubmed=7494039;		
RX	RA	Fukunaga M., Hamase A.;		
RA	RA	"Outer surface protein C gene sequence analysis of Borrelia		
RT	RT	burgdorferi sensu lato isolates from Japan.";		
RL	RL	J. Clin. Microbiol. 33:2415-2420(1995).		
RN	[5]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OG	Plasmid 1p54.			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID:139;			
RN	[1]	SEQUENCE FROM N.A.		
RP	RC	STRAIN-ATCC 35210 / B31;		
RC	RC	MEDLINE=96065943; Pubmed=9403685;		
RX	RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	RA	LaHirra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RT	RT	Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,		
RL	RL	Peterson J., Kellnerge A.R., Gockenush J., Salzberg S., Hanson M.,		
RN	RN	van Vugt. R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
AC	DT	15-DEC-1998 (Rel. 37, Created)		
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	OUTER SURFACE PROTEIN C PRECURSOR (PC).			
GN	OSPC OR BB619.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			

```

RA RA Garland T., Matthey L., McDonald L., Attiach P., Bowman C.,
RA Gairland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: X69596; CAA49306.1; -
DB EMBL: 001894; AAA16058.1; -
DB EMBL: D49497; BAA08457.1; -
DB EMBL: AE000792; AAC66329.1; -
DB TIGR: BB819; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW Complete proteome.
FT SIGNAL 1 18 BY SIMILARITY.
FT CHAIN 19 210 OUTER SURFACE PROTEIN C.
FT FT 19 N-ACETL DIGLYCERIDE (BY SIMILARITY).
FT LIPID 19
SQ SEQUENCE 210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match 33.7%; Score 928; DB 1; Length 210;
Best Local Similarity 98.4%; Pred. No. 5e-33;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MACNSGKDGNTSANSADSVGPMUTELISKITDSMAVLLAVKEVALLSIDEIAAKA 60
DB 17 ISCNNSGKDGNTSANSADSVGPMUTELISKITDSNAVLLAVKEVALLSIDEIAAKA 76
OY 61 IGKRIHONNGLDTEVENHNSGLLAGAAYAISTLIKOKLDGKNGEKLKEKIDAAKCKSEPTN 120
DB 77 IGKRIHONNGLDTEVENHNSGLLAGAAYAISTLIKOKLDGKNGEKLKEKIDAAKCKSEPTN 136
DB 121 KIKEKH7DLGEGVTDADAKAEAILKTNGCTKTGAEEIKGLFESVYLSKAAKEMIANSVK 180
DB 137 KIKEKH7DLGEGVTDADAKAEAILKTNGCTKTGAEEIKGLFESVYLSKAAKEMIANSVK 196
OY 181 ELTSPVAESP 191
DB 197 ELTSPVAESP 207

RESULT 2
OSC2_BORBU STANDARD; PRT; 212 AA.
AC Q08137;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (PC).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochaete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=PRO;
RX MEDLINE=92219995; Pubmed=1560779;
RA Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

```

```

RA Soutschek E.;
RA "Molecular analysis and expression of a Borrelia burgdorferi gene
RA encoding a 22 kDa protein (Pc) in Escherichia coli.",;
RA Mol. Microbiol. 6:503-509(1992).
RN [2]
RP SEQUENCE OF 1-205 FROM N.A.
RC STRAIN=DK26;
RA Theisen M., Frederiksen B., Ijebach A.M., Yunst J., Hansen K.;
RA Medline=94075528; PubMed=8253951;
RT "Polymorphism in ospC gene of Borrelia burgdorferi and
RT immunoreactivity of OspC protein: Implications for taxonomy and for
RT use of OspC protein as a diagnostic antigen.",;
RL J. Clin. Microbiol. 31:2570-2576(1993).
RC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or\_send\_an\_email\_to\_license@sib-sib.ch).
CC -----
DR EMBL; X62162; CAA44093.1; -.
DR EMBL; X73624; CAA52003.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR Prodom; PD001149; Lipoprotein_6; 1.
DR KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT SIGNAL 1 18
FT CHAIN 19 212 OUTER SURFACE PROTEIN C.
FT LIPID 19 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE 212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;
Query Match 32.2%; Score 886; DB 1; Length 212;
Best Local Similarity 90.7%; Pred. No. 2,9e-31;
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;
QY 195 SNSGCGDSASTNPADESAKGNLITEISKITSDSAFYLAVKEVETLVLSIDELAKKAIG 254
Db :|||||
20 NNSGCGDSASTNPADESAKGNLITEISKITSDSAFYLAVKEVETLVLSIDELAKKAIG 79
QY 255 QKIDNNNGLAALNNONGSLAGAYVISTLITEKLSKLNLEELTEIAKAKKCEEFNTK 314
Db :|||||
80 QKIDNNNGLAALNNONGSLAGAYVISTLITEKLSKLNLEELTEIAKAKKCEEFNTK 139
QY 315 LKSGADGKGDADDDHAKKAILKTHATTDGAKEFQDFSVGGLLAKAQAVALTNSKE 374
Db 140 LKSGADGKGDADDDHAKKAILKTHATTDGAKEFQDFSVGGLLAKAQAVALTNSKE 199
QY 375 LGHRNNSGDSASTNP--DESAKGP 397
Db 200 L-----TSPVVAESPGRP 212
RESULT 3
VM24 BORHE
ID VM24_BORHE STANDARD; PRT; 214 AA.
AC P33778;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24. PRECURSOR.
GN VMP24.
OS Borrelia hermsli.
OG Borrelia.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_taxid=140;
[1]
SEQUENCE FROM N.A.
RP

```

RC STRAIN-SSP, HSI SEROTYPE 24;
 RA MEDLINE-93133110; PubMed-1484486;
 RT Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
 "Subtelomeric expression regions of *Borrelia hermslii* linear plasmids
 are highly polymorphic.";
 RL Mol. Microbiol. 6:3299-3311(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L04786; AAA22964.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Signal; Plasmid.
 FT SIGNAL 1 18 PROBABLE MAJOR OUTER MEMBRANE
 FT CHAIN 19 214 VARIABLE MAJOR OUTER MEMBRANE
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 214 AA; 22541 MW; F1583F510246F7C7 CRC64;

Query Match 14.6%; Score 402.5; DB 1; Length 214;
 Best Local Similarity 44.9%; Pred. No. 4.9e-11;
 Matches 92; Conservative 37; Mismatches 59; Indels 17; Gaps 5;

OY 1 MACNNSGKNGTNSANSADSVKGP-----NLTEISKRTITSDNVLAVKEVALLSIDEL 56
 DB 17 MSCNNGGPE-----LKSDEVAKSDGTVDLAKYSKKIKKASFAASVKEVTLVKSDEL 71
 OY 57 AKAIGKRIKIHONGDTEYNHNSLAGAYASTLIKOKL-----DGLKNGLEKKEIDA 110
 DB 72 -AKAIGKRIKIHONGDTEYNHNSLAGAYASTLIKOKL-----DGLKNGLEKKEIDA 110
 OY 111 AKKCEFTFNKLEKHTDGLGEGVTDADAKKALIKTNGTKTGAELGLFESVEVLSKA 170
 DB 130 VSKSAEAFYQVKSHTDGLGEGVTDADAKKALIKTNGTKTGAELGLFESVEVLSKA 170
 OY 171 AKEMLANSVKELTSPVAVSPAMGS 195
 DB 190 AGEVEBAIKELTAPVAKKPSQNN 214

RESULT 4
 VMO3 BORNE STANDARD: PRT: 215 AA.
 ID VMO3 BORNE STANDARD: PRT: 215 AA.
 AC 002448; 11
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
 GN VMP3.
 OS *Borrelia hermslii*.
 OC Plasmid.
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.
 OX NCBI_TaxID=140;
 RN SEQUENCE FROM N.A.
 RC STRAIN-SSP, HSI SEROTYPE 3;
 RA MEDLINE-93133110; PubMed-1484486;
 RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
 RT "Subtelomeric expression regions of *Borrelia hermslii* linear plasmids

RT are highly polymorphic.";
 RL Mol. Microbiol. 6:3299-3311(1992).
 CC -1- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
 FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
 ANCHOR.
 CC -1- SIMILARITY: STRONG, TO VMP24.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L04789; AAA22967.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Signal; Plasmid.
 FT SIGNAL 1 18 PROBABLE MAJOR OUTER MEMBRANE
 FT CHAIN 19 215 VARIABLE MAJOR OUTER MEMBRANE
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 215 AA; 23139 MW; 684C7AD35F87C771 CRC64;

Query Match 14.5%; Score 398; DB 1; Length 215;
 Best Local Similarity 45.4%; Pred. No. 7.5e-11;
 Matches 93; Conservative 35; Mismatches 61; Indels 16; Gaps 5;

OY 1 MACNNSGKNGTNSANSADSVKGP-----NLTEISKRTITSDNVLAVKEVALLSIDEL 56
 DB 17 MSCNNGGPE-----LKSDEVAKSDGTVDLAKYSKKIKKASFAASVKEVTLVKSDEL 71
 OY 57 AKAIGKRIKIHONGDTEYNHNSLAGAYASTLIKOKL-----DGLKNGLEKKEIDA 110
 DB 72 -AKAIGKRIKIHONGDTEYNHNSLAGAYASTLIKOKL-----DGLKNGLEKKEIDA 110
 OY 111 AKKCEFTFNKLEKHTDGLGEGVTDADAKKALIKTNGTKTGAELGLFESVEVLSKA 170
 DB 131 VKKSEAFYQVKSHTDGLGEGVTDADAKKALIKTNGTKTGAELGLFESVEVLSKA 170
 OY 171 AKEMLANSVKELTSPVAVSPAMGS 195
 DB 191 ANDAVETVIKELTAPVAKKPSQNN 215

RESULT 5
 YD86_SCHPO STANDARD: PRT: 1957 AA.
 ID YD86_SCHPO STANDARD: PRT: 1957 AA.
 AC 010411;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
 GN SPAC1F3.06C.
 OS *Schistosoma* *schistosomae* pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: 270690; CAA94624.1; -
KW Hypothetical protein.
SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 7.7%; Score 210.5; DB 1; Length 1957;
Best Local Similarity 22.4%; Pred. No. 0.065;
Matches 157; Conservative 116; Mismatches 218; Indels 209; Gaps 35;

```

OY 8 KQGNSTANSADG---SVKGNPLTEISKITDSDNAVLAV--KVEALLSIDELAIAKIG 63
DB 360 KDSRTNSQLEEMVELKSNRT--HSQULTDAESKISSEQENKSLKSGSIDE----- 410
DB 64 KIHONN-----GLDTEYHNHNSLAGAYAISTLIKOKLDGLKNEGLK--EKIDA-A 111
DB 411 --YONNLSSKDKMVKVQSSQLEARSLSLAHATGKLAIEINSEED--FONKKIKDFEKIEDL 467
OY 112 KRCSEFTFKLKEKHTDLDKEGVTDADAEALTKNGTKTKAEELGKLFESV--EVLSK 169
DB 468 RACLNSSSMELKEKSLIDKQDELNLNREQIKE---OKKVESTQSSLSQLODILNE 523
OY 170 AAK-EMLANSVKELTSPVVAESPAMGSGSGDASTNPADSAKGNPLTEISK--- 224
DB 524 KKKHEYESQELNELKGELOTEI---SNSEHLSQSLTSLAEKEAAVATNNELSSKSL 579
OY 225 IYDSNAF-----VLAKVEVTLVLSIDELAKK--AIGOKIDNNNGL----- 263
DB 580 QTLCAHFOBKLAKSVMQLEKNEONFSSLDTSFKKLNESHQELNNHQITQOLKDTSSKL 639
OY 264 -----AALNNOGSLGAGAYAISTLITEKSL----- 291
DB 640 QOLLEKLANFEQKESTLSDENNDL-----RTLLKLESNKSLIKQEDVDLSLE 688
OY 292 KNELEKLEIAKAK-----KCESEFTNKLKSGHADLQKQ-----DATDDHAKAA 335
DB 669 KNIQTLKEDLKRSEALFKSLKLEKNLEKREVINDLKGKHEHLEAQRNDLHSLSDAKNTNA 748
OY 336 ILKTHATADKAKFEKDFEFSEVEL-----LKAQVALTNSVVELGH----- 377
DB 749 ILSSELT--KSEDEVKRLTANVELTLDQSKAMKOSFTSLVNSYSISNLVYHELDRDHVM 806
OY 378 --RNNS--GGDSASTNPDESAKGNPLTVISKITDSDNAFLAVKEV--EALLSIDELSK 431
DB 807 QSONNTLSESEKLTDCENLTQNMTLI-----DNVQKLMHKNHVNQESKVELKEVNG 860
OY 432 AIGKKIKN-DOGLDNEANRNSLIAGAVEISK---LITOKLSVNS----- 473
DB 861 KLSLDLKNLSSLVNVAISDNDQILTQLAELSKNYSLEQSAQSLSLKSLAEKQLLHT 920
OY 474 --EEL-----KKRIKEAK--DCSQKFTTKLSDSHAEIGIOSVODDD--AKKALIKT 518
DB 921 ENEELHRLDKLTGKLTIESKSSDLGKTLTAQOE-----ISNLKEEMASOALITSV 974
OY 519 HGTDRKAKLELELFKSLSEKRAQAALTNVSKELTNPV 558
DB 975 KSKLD-----ETLSKSSK--LEADIEHLKNKV 999

```

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RN J. Cell Biol. 113:245-260(1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McCellan M., Tao N.,
RA Kendrick K.E.;
RN Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RN Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
CC [3]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Query Match 7.5%; Score 207.5; DB 1; Length 1790;
Best Local Similarity 21.4%; Pred. No. 0.079;
Matches 143; Conservative 104; Mismatches 239; Indels 161; Gaps 22;

```

OY 3 CUNSGKDG-----NTSANSADSVKGNPLTEISKITDSDNAVLAVKEVALSS 52
DB 929 CUNLSKEKEHISKELVEYKSRQSHDNLV--AKLTKKLSLANNYKMQA--ENESLIKA 984
OY 53 IDEIAKAIKGIHONNGLD--TEYNHNSLAGAY-----AISTL-----IKOK 95

```

```

Db 985 VESKNESSIQLSNQKIDMSOEKENFOIERGSIERNIEOLKKTITSDLEQTEITISK 1044
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 96 LDGLRNEG-----LKEIDAAKCSFTNKLKE-----KHTDLGKEGVTD 136
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1045 SSSKDEYESSQISLLEKLEKLETTATANDENVNKISLETKRLEELAEALAYKNLKNLETK 1104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 137 AAKKEILK-----TNGTKTK-----GAELGKLFESVEVLASKAA 171
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1105 LETSEKALKEVENEHELEKEKIOLEKEATEKQOLNSLRANLESLEKEHEDLAQLKRY 1164
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 172 KEMLAN-----SVKELTSPVAASPMAGSNGKGDASIN 207
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1165 EOJANKKEQVBEISQLNDEITSTOOENESITKKNDLEGEVYKAKSTS---EEOGNLK 1221
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 208 PADESAGPNLTETISKKTIDSNAPVL-AVKEVETLVLSIDEL-----AKKAIGOKIDNN 260
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1222 KSEIDALNTQIKELKKKNETNEASLESIKSVSEETVKIKELQDECNFKKEVSELEDKL 1281
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 261 NGLAALNONGSLAGAVASTLITELKSLKLNLEELKEIKAKKCSFTNKLKSGHA 320
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 1282 KASEDKNSKYLLEKKESEKIKELDKITELKIOLEKITNLKAKKSESELSRLK---- 1337
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 321 DLGKODATDHAKAATLKTHATTDGAKKFKDLFESVEGLTAAQVALTNSVKEIGHNN 380
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1338 -----KT-----SSEKNAEEOLEKLEKNEIQKNQAFERKLLN- 1373
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 381 SGDSASTNPDESANGPNLTIVISKKTIDSNAPFLAVKEVALLSIDELSKAIGKIKND 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1374 ---EGSSTTIOEYSEKIN-TLEDELIRLQNEENELKAKELIDNRESELEKVS-----LSND 1423
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 441 GTLDNEANNESELIGAVTISKLTIOKLSVLS-----ELKKKIKKEAKDCQKPT 491
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1424 ELLEEQNTIKSLQDEILSYKDKITRNDKLLSTERNKRDLESLEKQRAQESAKAYE 1483
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 492 TIKDSDHAELGSIQVODNNAKKAITLHTGTDKAKEL-----ELFKSLESLSKAAQAL 547
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1484 EBLKLEESSSEKAELEKSEKMKLESTIESNETELSSMETIKKSDPEKLEQSKSA- 1542
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 548 TNSVKEL 554
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1543 EEDIKNL 1549
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

CC entties requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M88097; AAA29743.1; -.
DR HSSP: P36956; 1AM9.
KW Malaria; Receptor; signal; Transmembrane.
FT SIGNAL 1 17
FT CHAIN 18 2869
FT DOMAIN 18 2807
FT TRANSMEM 2808 2826
FT DOMAIN 2827 2869
FT SITE 1030 1032
FT SITE 2599 2601
SQ SEQUENCE 2869 AA; 330213 MW; B9DBE442205BECF CRC64;

Query Match 7.5%; Score 207.5; DB 1; Length 2869;
Best Local Similarity 20.4%; Pred. No. 0.13;
Matches 127; Conservative 116; Mismatches 260; Indels 121; Gaps 24;

Oy 25 NLTETSKIT-DSNAVLAVKEVALLSIDE-----IAAAGKRIHONG----LDT 73
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1141 HINTVAHGITSNNKNEILKSVKEVEDKLNEQNEDEYKKVKNPENKOLBAIGSMSKLE 1200
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 74 EYNHNSILAGAVASTLTKOKLDGLKNEGLKEIDAAKCSFTNKLKEKHTDLGREG 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1201 VINKHVESEMTQLESTANTLKSNAKGENEHDELELTKTGQMRDIYELKLAEL-KEG 1259
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 134 VTDADAKAAILKTNGTKTGAEE-LGKLFESVEV-----LSKAAKEMANSVELETPVA 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1260 TVN-ELKADANEKANKVEPPERNIIGHVLERITVEKDKAGKAYEEN--NSLTKIKELQ 1316
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 189 ESPAMGNSGKGDSASTNPDESAGKPNLTETISKKTIDSNAPFLAVKEVETLVLSIDEL 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1317 ET-----SDDSQNELVTTSTKHLNENAKGYEDVKKNEEDSIQREK 1358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 249 AKKAIGOKIDNNNGLAALNONGSILAGAVASTLITELKSKLNLEL-----KTEIA 302
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1359 AKSL-ETLDEKMKLVQOVNMM--LQSAIQAGAGISKELENEKGVIELLISTNNSILE 1413
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 303 KAKKCSFTNKLKSGHADLGKODATDHAKAAI-----LKTHTATDKAKKFKDLFESV 357
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1414 YKKNSESVRSQLANGFTAKEGEKNASARLAELKQIVDLDYSDIDDKVKKI 1473
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 358 EG-----LKAQVALT-----NSVKEL--GHRNN-----SGDSASTNPDESA 394
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1474 EGKREILKMKESALFWESESEKFKOMCSSHMEKAKKIEYLKNNGDGKANTDSQ 1533
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 395 KGNPLTVISK-----KITDSNAFLAVKEV-----EALLSSI-----D 427
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1534 MEKVGYVSKAEHAFHTVBAQVDTKAFCESTIYAAYTKMDNLFENSLMEVKVCKEKKND 1593
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 428 ELKSAIGKIKINDGTLDNEANRNESLIAGAVETSKLTIOKLSVJANEELKKIKKEAKDS 487
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1594 EAEKYSAKLKPYDGRKAKAVSENERKISLEKAKYKKESSOLN-DVSTKSLQIDNOR 1652
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 488 QKFTTKLKD-----SHAELGISOV---QDNNAKKAITLHTGTRDKAKKELEEL 532
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1653 QQLDSVLSNIGRYKQNALQFDSADRSKMSVLPISLSELGAKESIDKVKAKSEKYLETV 1712
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 533 FKSLSELSKAAQALTNVSKELTN 556
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1713 QNEMSRIN-VEEGSLTIDDKKTTD 1735
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
RBP1_PLAVB STANDARD: PRT: 2869 AA.
AC 000798;
DT 01-APR-1993 (rel. 25, Last sequence update)
01-APR-1993 (rel. 25, Last sequence update)
01-OCT-1996 (rel. 34, Last annotation update)
RETILOCYTE BINDING PROTEIN 1 PRECURSOR.
GN RBP1.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites."
RL Cell 69:1213-1226(1992).
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
HUMAN RETICULOCYTE CELLS.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.

```

```

RESULT 8
MYSN_ACACA STANDARD: PRT: 1509 AA.
AC P05659;
DT 01-NOV-1988 (rel. 09, Created)
01-NOV-1988 (rel. 09, Last sequence update)
01-NOV-1988 (rel. 09, Last sequence update)
20-AUG-2001 (rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.

```

OS Acanthamoeba castellanii (Amoebozoa).
OX Eukaryota; Acanthamoebidae; Acanthamoeba.
NCBI_taxid=5735;
RN
RA SEQUENCE FROM N.A.
RX MEDLINE:8-3083395; PubMed:3040773;
RA Hammer J.A., Iit, Bowers B., Paterson B.M., Korn E.D.;
RT *Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail.*
RL Cell Biol. 105:913-925(1987).
CC
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TALPIECE WHERE
CC THE REGULATOR PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMAIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: Y00624; CA68663.1; .
DR PIR: A27224; A27224.
DR HSP: P08799; IMND.
DR InterPro: IPRO00048; IQ.
DR InterPro: IPRO01609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR PRINTS: PR00193; MYOSINHENV.
DR ProDom: PD00035; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
DR KEGG: Metabolism; Colled coil; Actin-binding; ATP-binding; Calmodulin-binding;
DR KEGG: Myosin; Colled coil; Actin-binding; Phosphorylation; Multigene family.
FT METYLATION; Alkylation; Phosphorylation; Methylgene family.
FT DOMAIN 1 789
FT 790 819
FT 848 1509
FT 848 1226
FT DOMAIN 1227 1252
FT 1253 1509
FT DOMAIN 1253 1482
FT 1483 1509
FT NP_BIND 182 189
FT 660 682
FT DOMAIN 766 780
FT 133 133
FT MOD_RES 700 700
FT MOD_RES 1489 1489
FT MOD_RES 1494 1494
FT MOD_RES 1499 1499
FT MOD_RES 1509 AA: L17201 MW: ZCE49BE51173D1017E CRC64;
SO SEQUENCE

Db	934	ELQEFSTASNDILLEQKREKLEAEKGEKLEKASLEEEERNNKALQEQAKTKVESERNELQDEYD	993
QY	78	NSGLGAVYAIISTLKIKLDGLKNGEKKETIDA---AKKSEFTNTLKKRHIDLGEGV	134
Db	994	EA-----AAHSIKKKEEDLSRE-LRETKDALADAMENISEILRSL--KNTERGADVV	1043
QY	135	TDA--DAKEAIIKLTNGTKTKGAEBELCKFESEVLSKAAKEMANSVKELTSPVYAESPA	192
Db	1044	RNEIIDLVTATKLTQLETKTKKSLKEELAQTRQALEE-EKSGKEAASAKAQIGOOLEPARSE	1102
QY	193	MGNSNGCGGDSASTNPADDESAKGNPULEISKTTDSNAFLAVKAEVETVLISDELAKKA	252
Db	1103	VDSLKSK---LSAAEKSLKTAQDN-RDDEQOLEDE-----RTYRANADK-QRKA	1147
QY	253	IGQKI-DNNNGLAALNNQNSGLAGAYAIISTLTTEKLSKLNLEELKTEYAKKGCSEEP	311
Db	1148	LEAKLTLEDEQVYALDQKMAAAQAQKTLTKYQDETKRRLEEADEASAARLEKERRNALDE	1207
QY	312	TNKLKS---GHADLGQODATDHDAAKAIL-----KTHATTDKCAKEPKDFESVEGL	360
Db	1208	VADLTLDLDAERSGAQOQRKRLNTRISELQSELENAKPTGGASSEVEKRLLEGELERLEE	1267
QY	361	LKAQAAYALTNVSELELHRN-----NSGGSASTNPDESAGCNLYVISKRI-TDSNAFL	413
Db	1268	LITQAERARAAAEKLNKDLKANLELEELNQEDDARDADNDKLVKD-----NRKLKADIDDEAR	1321
QY	414	LAVKEVE-----ALLSIDELSKAIGKKIKN-DGTLDNEN--RNSLTAGAV	458
Db	1322	IQLEEQDAKSHADSSRRLLAEIEELKRVAKETSDKQAKOQKANYORENSLSKA---	1378
QY	459	EISKLITQKLSVNSELELKKIKKEADCQKFTTKKIDSHAEIGIOSYODNKAQKILTKT	518
Db	1379	-----DROSIERNRNDAE-----RQYRDLRAQDLDALSRDSEKRA-----	1414
QY	519	HGTQDKGAKLEBELFK-----SLESTISK	541
Db	1415	---KEKSVANRELKKKVVLDREROSLSELSK	1442
RESULT	9		
MYSN_DROME			
ID	MYSN_DROME	STANDARD:	PRT: 2017 AA.
AC	09323:		
DT	01-JUN-1994	(Rel. 29, Created)	
DT	01-JUN-1994	(Rel. 29, Last sequence update)	
DT	20-AUG-2001	(Rel. 40, Last annotation update)	
DE	MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).		
GN	ZIP.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	Ncbi_Taxid=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Embryo.		
RX	MEDLINE=90349606; PubMed=2117279;		
RA	Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;		
RT	"Complete sequence of the Drosophila nonmuscle myosin heavy-chain		
RT	transcript: conserved sequences in the myosin tail and differential		
RT	splicing in the 5' untranslated sequence."		
RL	Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320.(1990).		
CC	-!- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR		
CC	CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.		
CC	-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A		
CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.		
CC	-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.		
CC	-!- SIMILARITY: CONTAINS 1 IQ DOMAIN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M35012; AAA28713.1; .
DR PIR: A36014; A36014.
DR PIR: B36014; B36014.
DR HSP: P08799; IAMD.
DR Flybase; FBgn0005634; zip.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; Myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
Myosin; Alternative splicing; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding.

FT DOMAIN 1 829 MYOSIN HEAD-LIKE.
FT DOMAIN 830 859 IQ.
FT DOMAIN 886 2017 COILED COIL (POTENTIAL).
FT NP_BIND 225 232 ATP.
FT DOMAIN 250 260 25 KDA/50 KDA JUNCTION.
FT DOMAIN 682 694 50 KDA/20 KDA JUNCTION.
FT DOMAIN 705 727 ACTIN-BINDING.
FT DOMAIN 742 758 REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT DOMAIN 1303 2017 LIGHT MEROMYOSIN (LMM).
FT DOMAIN 1303 1970 ALPHA-HELICAL TAILPIECE.
FT DOMAIN 1971 2017 GLOBULAR TAILPIECE.
FT VAAPLIC 1 45 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 2017 AA; 232016 MW; 73E3CB02BBA8F2528 CRC64;

Query Match 7.0%; Score 193.5; DB 1; Length 2017;
Best Local Similarity 23.3%; Pred. No. 0.35;
Matches 131; Conservative 87; Mismatches 181; Indels 163; Gaps 27;

QY 26 LFEISKKITDSNAVL-----AKVEYALLSID-----ELAKAIGKRIHONGDTEVN 76.
DB 1136 LIQTLRIDESAATRAQAQAKRELESQAEIQEDLEAKARAKAREKRRD----- 1187
QY 77 HNGSLLAGAYAISTLIKOKLDGLKNGLEKIDK-----AKKSEFTTKLKEKHTD 128
DB 1188 -----LSELELAKNE-LIDSIDTTAAQDELKSKREQELATLKKSLEET 1231
QY 129 LKEGVTDADAKAELIKTNGTKGAEEELGKLFESVEYLSKAKEMLANSVKELTSPVVA 188
DB 1232 VVHECVL-ADMMH-----KHSQELNSINDOLENKR-AKTVL----- 1266
QY 189 ESPAMGNSGKGGDSASNPADESAAGPVLTEISKITDSNAFVLAVK--EYETVLSTD 246
DB 1267 -EKAGTLEAEVADLATELRSVNSROEN--DRRRKQSQAEIQLVETLAEIERARSEQ 1323
QY 247 ELAKKAIGKIDNNNGLALNNQ--NGSLLAGAYAISTLITEKLSKLKLELKEKTIKA 304
DB 1324 EKCCK-LQOEAEEN-----ITNQLAEELKASA-----AVKSASNNESOLTER 1364
QY 305 KKCSEFT-----NKLKSGHADLGKODATDHAH-----AALIKTHATTDKGA 347
DB 1366 QQLLEETNQKLGSLSKLQIGSEKALQEQLEEDDEAKRNERKLAETVTTQOELIKKA 1424
QY 348 KEFKDLFESVEGLKAAQ--VALINSVKELGRNNSGDSASTNPDESAGKGNLTIVISK 404
DB 1425 EEDADLAKLEEGKKRLNKDIEALEROVKELIAOND-----RLDKSKK-----KIOS 1471
QY 405 KITDSNAFLA-----VKVEYALLSIDEL--SKAIGKKIKND-GTLDNEARNNESLING 456
DB 1472 ELEDATILEAORTVLELEKKRKNPKDLAEKALISEQIAOERDTAREAREKE----- 1526

QY 457 AVEISKLITOKLSVNSLEELKKIKAEKDCSQFTTKLKDSHAELGICSVODNNAKATL 516
DB 1527 -----TKVLSTV--SRLEDAEPDKIEDLENKRRT-----LQNELD-----LA 1561
QY 517 KTHGTRDKGAKLEELFKSLES 538
DB 1562 NTQGTADKRVNHELEKAKRALES 1583

RESULT 10
MS2_DICDI STANDARD; PRT: 2116 AA.
AC P08799; IAMD.
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN II HEAVY CHAIN, NON MUSCLE.
GN MHCA.
OS Dictyostelium discoidium (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de Lozanne A., Weinand L.A., Spudich J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoidium."
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN [2]
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN-AX2;
RX MEDLINE=90353583; PubMed=2387408;
RA Lueck-Vietmeter D., Schleicher M., Grabatin B., Wipplert J., Gerisch G.;
RT "Replacement of threonine residues by serine and alanine in a phosphorylatable heavy chain fragment of Dictyostelium myosin II."
RL FEBS Lett. 269:239-243(1990).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Wagle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the Dictyostelium myosin heavy chain."
RL FEBS Lett. 227:71-75(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Holden H.M., Rayment I.;
RT "X-ray structures of the myosin motor domain of Dictyostelium discoidium complexed with MgADP, BeFx and MgADP.Alfa-".
RL Biochemistry 34:8960-8972(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the truncated head of Dictyostelium discoidium myosin to 2.7-A resolution."
RL Biochemistry 34:8973-8981(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611550;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the Dictyostelium discoidium myosin motor domain to 1.9-A resolution."
RL Biochemistry 35:5404-5417(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX MEDLINE=97452580; PubMed=9305951;
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammaS, and MgAMPNP complexes

RT of the Dictyostelium discoidium myosin motor domain.";
 RL Biochemistry 36:11619-11628(1997).
 RA X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RP MEDLINE-98070605; PubMed-9405148;
 RA Baer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
 RT "X-ray crystal structure and solution fluorescence characterization
 of Mg²⁺(3')-(O-(N-methylanthraniloyl) nucleotides bound to the
 Dictyostelium discoidium myosin motor domain.";
 RL J. Mol. Biol. 274:394-407(1997).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (HMC), 2 ALKALI
 LIGHT CHAIN SUBUNITS (LMC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 (MLC-2).
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
 CORTEX.
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN BE FURTHER
 SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
 THE ACTIN-ACTIVATED ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
 ACTIVITY. PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
 POSITION (688).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14628; AAA33227.1; -
 DR PIR: A26655; A26655
 DR PIR: S00250; S00250
 DR PDB: 1MAA; 03-DEC-97
 DR PDB: 1AMD; 17-AUG-96
 DR PDB: 1MMG; 03-DEC-97
 DR PDB: 1MMN; 03-DEC-97
 DR PDB: 1MND; 17-AUG-96
 DR PDB: 1MNE; 17-AUG-96
 DR PDB: 1VOM; 23-DEC-96
 DR PDB: 1LVK; 28-JAN-98
 DR Dictydb: DD01008; mhca.
 DR InterPro: IPR000408; IO.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS0096; IQ; 1.
 KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
 KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
 FT DOMAIN 1 761
 FT DOMAIN 762 791
 FT NE_BIND 817 2116
 FT NE_BIND 179 186
 FT DOMAIN 638 660
 FT DOMAIN 738 752
 FT MOD_RES 130 130
 FT MOD_RES 678 678
 FT MOD_RES 1823 1823
 PHOSPHORYLATION (SH-1).
 PHOSPHORYLATION (BY MHCK).

FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
 FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
 SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;
 Query Match 7.0%; Score 193.5; DB 1; Length 2116;
 Best Local Similarity 20.8%; Pred. No. 0.37;
 Matches 140; Conservative 92; Mismatches 257; Indels 185; Gaps 24;
 QY 9 DGNTSANSADSVGK-----PNLTETSKKTTDSNAVLAVKEVEALLSIDETAAKIG 62
 DB 964 DGSDPTISRLRKIDELQVEEELTESPESKRGVL--EKTRVRLQSELDLTVR--- 1018
 QY 63 KKHQNNLDQFEYHNGSLGAVASIFLIQKQDGLKNEGLKEKIDAKKCEFTNKL 122
 DB 1019 -----LDSETDKSLLQKKKLEBEILOVOQALAE-----TAAKLAQENANKKL 1064
 QY 123 KEKHTDQKEGVTDADAKAELIKTNGT-----KTKGA----- 154
 DB 1065 QGEYTELNEKFNSEVTLANSVEKSKKTLQSLQVAVNNELDEKKNRDALEKKKALDAML 1124
 QY 155 EELGKLPES-----VEYLSKAKAKEMANSYKELTSFV----- 187
 DB 1125 EEMKDQLESTGEGKSLYDLKVKQSDMEALRNQISELOSTIAKLEKIKSTLEGEVARLQ 1184
 QY 188 ----AESFAMSGNGKGG-----DSASTNPDESAGPNLTETSKKTTDSNAFVLAVK 236
 DB 1185 GELAEADQAK-SNEKQKKKVEYLDLQKSLQALAEETAKQALDKKKLEBE-----LS 1237
 QY 237 EVELVLVSLDELAKKAIQKIDNNGLAALNNQNSLAGAVASTLTTELK---SKLKN 293
 DB 1238 EVQT---OLSEANNKNVNSDSTNKHLETFSPNNLLELEAEQAKAQALEKKRLGLESELIK 1294
 QY 294 LEE-----LKEIIRAKKCKSEFTNKLKSGHADLGKOATPDHAKAA 335
 DB 1295 VNEDELEKKKOKESNEKRYDLEKEVSELDQITEEYASKAAYEAKNKKSELDETIKRQ 1354
 QY 336 ILKHTATTDKAKKEFKDL-----FESVEGL-----KAAQVALTNSYKEIGHR 378
 DB 1355 YADVVSRSKQVEQLKTLQAKNEELRMTAEAEQDLRAERSKKKAEDELEAVKNLEEE 1414
 QY 379 -----NNSGGSASINPD-ESAKGN-----LTVISKKTIDS 409
 DB 1415 TAKRVKAEKAMKAKETRYRSKSELDAKNVSSQYQIKRLNELSELRLAEADERC 1474
 QY 410 NAFLLAVEVEFALSSI-DELSKAIQKKIKNDGFLDNEANNEISLAGAVISKLITQKL 468
 DB 1475 NSAIKAKKTAEALSLESLKDEIDANNNAKAKAE-----KRSKELEVRVALESESLDKS 1527
 QY 469 SVLNSEELKKRIKAEKDCSOKFTTKLKDASHAELGIQSVQD-DNAKKAILKTHGTDKGAK 527
 DB 1528 GTVAVFERKKDAIDDLRLARL-----DRETESIKRSDDEKKNRKQFADLEAKVEEAKR 1582
 QY 528 E---LEELFKSLES 538
 DB 1583 EYVTDRLKRLKLES 1596
 RESULT 11
 REST_CHICK STANDARD; PRT; 1433 AA.
 ID REST_CHICK
 AC 042184; 042228; 057563; 057564;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (Clp-170).
 GN RSN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]


```

RP SEQUENCE FROM N.A.
RX MEDLINE-98137792; PubMed-9469933;
RA Griparic L., Volosky J.M., Keller T.C. III;
RT "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RL Gene 206:195-208(1998).
[2]
RN
RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC TISSUE-Pectoralis muscle;
RA Griparic L., Keller T.C. III;
RT "Identification and expression of two novel CLIP-170/Restin isoforms
expressed predominantly in muscle.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
THAT LINKS ENDOCYTTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CYTOSKELETON (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF014012; AAC60344.1; -
DR EMBL; AF020764; AAC60345.1; -
DR EMBL; AF045650; AAC03547.1; -
DR EMBL; AF045651; AAC03548.1; -
DR InterPro; IPR000938; CAP-GLY.
DR InterPro; IPR001878; ZnF_CCHC.
DR Pfam; PF01302; CAP-GLY; 2.
DR SMART; SM00343; ZnF_C2HC; 1.
DR PROSITE; PS00845; CAP_GLY_1; 2.
KW Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT FT DOMAIN 79 121 CAP-GLY 1.
FT FT DOMAIN 144 207 SER-RICH.
FT FT DOMAIN 235 277 CAP-GLY 2.
FT FT DOMAIN 305 332 SER-RICH.
FT FT DOMAIN 351 1353 COILED COIL (POTENTIAL).
FT FT DOMAIN 1414 1427 CCHC-BOX.
FT FT VARSPLIC 458 492 MISSING (IN SHORT ISOFORM).
FT FT VARSPLIC 458 492 TOTLENHARIELSGLFETKADLORELEDR ->
CONFLICT 309 309 RKR0ISEDPEN (IN ISOFORM CLIP-170(11)).
CONFLICT 440 440 S -> GGGSKYS (IN ISOFORM CLIP-170(11)).
CONFLICT 440 440 T -> RKR0ISEDPEN (IN ISOFORM CLIP-
SEQUENCE 1433 AA; 161026 MW; 5631CEB683498E23 CRC64;
7.0%; Score 193; DB 1; Length 1433;
Best Local Similarity 22.6%; Pred. No. 0.25;
Matches 137; Conservative 98; Mismatches 280; Indels 92; Gaps 23;
OY 19 ESVKPNLTETSKRTDTSNAVLAVKEVALLSIDETIAAKIGKKIHONNGLDTEY--- 75
DB 708 ESEVDQHLVEMEDTLNKLQLEAEIKYKELDVLQAKCNE-QTKLIGSLTQOIRASEEKLIDL 766
OY 76 -----NINNG-----SLLAGYVAISTLKOKLUDGLKNE--GLKER-IDAAKCSST 117
DB 767 AALOKANSEKLEIQLSEQLOAAEKQIONLETEVVSNTKELQKEKLLDLEKNLSA- 825
OY 118 FTKNKLK---EKHTDQKRGCVTDA-----DAKEAIIKTGTGTGKA---EEIGKLFEE 162
DB 826 -VNOVKDSELEKELQTLKEFTKTSADVCAENAOAOMETINKLNQKREOFALMSSELEQLKS 884
OY 163 SVEVLTKAAKEMLANVSELTSPVAASEPANGSNSGKGDSASTNPADS--AKGPNLTET 220

```

Dd	885	NLIEMERKIKERERREQULTEAVKLENDIAELIMSSGSSNAQLMKMDLELRLKREJQ	944
Qy	221	ISKRTIDSNAEVLAVK-----EETVLVSLIDELAKKAIGOKIDNNGLAALN	267
Dd	945	IQLTELTANERKAVQLOKNEVBOGTAKAEQSOQETLTKHOEEL-KKMDQLTDMKKOMETSQ	10030
Qy	268	NQNSSLAG-AVAISLTITREKLSKLNLELKEIKAKKCCSEFTNLKSGHADLQGD	326
Dd	1004	NQYKDLOAKYEKETSEMITHDADIGFKONLLDAEALKAOKKNDLETOAEELKRO-	10620
Qy	327	ATDDKAALIKTHTATDKCAFEKFDLFEVEGLLAAQVALTNSVKEIGHNNSGGSA	386
Dd	1063	--AEQAKA-----DKAAEE---VLOTMEVYTEKAKITHQEKITETLASENSRQNE	11080
Qy	387	STNPDESAGKNLTVISKRTIDSNAEVL-LAVKEVEALLSSIDELSKAIGKIKINDGTLDN	445
Dd	1109	KLQNELDMLKQNNLKNBEELTKSKELLNENKKVEELKKEFEALKLAAQSQQLAALQ	11680
Qy	446	E-ANRNESLAGAVEIS-KLITQKLSVLNSF-----ELKKIKKAKDCSGKFTTK	493
Dd	1169	ENVVLAELGSRDESVTSHQKLEERSVYLNQNLLEKKKRESTLKKKEIDERSLQK----	12250
Qy	494	LKDSHAELGQSVDDNKAAILTKTGTRKDKGKKELEELFKFSLEISAKAQAALTNSVKE	553
Dd	1226	ISDTSLITQDDELEKLRNEITVLNRG-ENASAKTLQSVKTLSE---DKLLEEKVKN	12800
Qy	554	LTPNPVA	560
Dd	1281	LEOKLKA	1287
RESULT 12			
MYSB_CABEEL			
ID	MYSB_CABEEL	STANDARD;	PRT; 1966 AA.
AC	P02566;		
Dt	21-JUL-1986	(Rel. 01, Created)	
Dt	21-JUL-1986	(Rel. 01, last sequence update)	
Dt	20-AUG-2001	(Rel. 40, last annotation update)	
DE	MYOSIN HEAVY CHAIN B (MHC B).		
GN	UNC-54 OR MYO-4.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;		
OC	Rhabditidae; Peloderiinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=63273600; PubMed=6576334;		
RA	Karn J., Brenner S., Barnett L.;		
RT	"Protein structural domains in the Caenorhabditis elegans unc-54		
RT	myosin heavy chain gene are not separated by introns.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).		
RN	[2]		
RP	SEQUENCE OF 850-1966 FROM N.A.		
RX	MEDLINE=82272395; PubMed=7202124;		
RA	McLachlan A.D., Karn J.;		
RT	"Periodic charge distributions in the myosin rod amino acid sequence		
RT	match cross-bridge spacings in muscle.";		
RL	Nature 299:226-231(1982).		
RN	[3]		
RP	SEQUENCE OF 1876-1966 FROM N.A.		
RX	MEDLINE=83232892; PubMed=6571695;		
RA	Wills N., Geeteland R.F., Karn J., Barnett L., Bolten S.,		
RA	Waterston R.H.;		
RT	"The genes sup-7 X and sup-5 III of C. elegans suppress amber		
RT	nonsense mutations via altered transfer RNA.";		
RL	Cell 33:575-583(1983).		
CC	-1- FUNCTION: MUSCLE CONTRACTION.		
CC	HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)		
CC	AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2)		
CC	SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.		
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING		
CC	CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPAAPPTIDES.		

```

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC - PTH: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC - MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC ELGANS.
CC - MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC - SIMILARITY: CONTRAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: J01050; AAA28124.1;
DR EMBL: V01494; CAA24738.1;
DR PIR: A02992; MMKW.
DR HSSP: P08799; IMND.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00063; myosin_head.
DR Pfam: PF01576; Myosin_tail.1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head.1.
DR SMART: SM00242; MYSC; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family.
FT DOMAIN 1 850
FT DOMAIN 851 1966 MYOSIN HEAD-LIKE.
FT DOMAIN 851 1164 COILED COIL (POTENTIAL).
FT DOMAIN 1165 1176 ALPHA-HELICAL TAILPIECE (S2).
FT DOMAIN 1165 1966 HINGE.
FT NP-BIND 177 184
FT NP-BIND 184 184 LIGHT MEROMYOSIN (LM).
FT DOMAIN 665 687 ATP (BY SIMILARITY).
FT DOMAIN 769 783 ACTIN-BINDING.
FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 705 705 METHYLATION (SH-1).
FT MOD_RES 715 715 ALKYLATION (SH-2).
FT MOD_RES 1337 1337 E -> R (IN REF. 2).
FT MOD_RES 1880 1880 I -> L (IN REF. 2).
FT CONFLICT 1880 1880
FT CONFLICT 1966 1966
FT SEQUENCE 1966 AA; 225125 MW; B6F0BB2FE27B67F CRC64;

Query Match 7.0%; Score 192; DB 1; Length 1966;
Best Local Similarity 21.5%; Pred. No. 0.39;
Matches 127; Conservative 104; Mismatches 213; Indels 146; Gaps 28;

OY 16 SADESVKPNLITEISKITDSNAVLAVK---EVEALLSSIDEI--AAKAIGKIHQ--NN 69
DB 1358 SLREEIEGKN--ELRLQSLKANDIDQWKAKEGEBELAA--DELEAKKRQOKINELQOE 1414
OY 70 GLDTEVNHNSL--LAGAYALSTILKOKLDGLKNEGLKEKIDAKKCSFTFNKLEKHTD 128
DB 1415 ALDAANSKASLEKTSRLVGLDDQAVVERANGVAASLEKKQKGFDKITIDEMRKKTDD 1474
OY 129 LGKEGVTDAKFAILKTNGT---KTKGAEGLKLEESVEVSKAKKEMLANSVKELTSP 185
DB 1475 LAEE---LDGAQRDLRTSTDLFAKNAOE--ELAEVEGELRRERKS--LSQEIIDLTDQ 1527
OY 186 VVAESPAMSGSGKGDASSTNPADASAGPNLITEISKITDSNAFVLAKEVETLVLSI 245
DB 1528 L-----GEGGRS-----VHMQKIIIRL 1545
OY 246 DELAKKAIGOKIDNNNGLAALNNKNGSLAGAYALSTILTEKLSIKNLEELKTEIAR-A 304
DB 1546 -EIEKEEIQHALDEAE--AALEAESKYLRAQVEVS-----QIRSEIERKRI 1588

```

```

OY 305 KKCSEFTNKLKSGHADLGQDATDDHAKA-----AIIKTHATTDGKAKKFKDLFESVEG 359
DB 1589 QKEKEEFNTK-----NARALESMQASLEIEA---KKAELIRKRLKLEG 1632
OY 360 LKKAQVALTNSVKEIGHRRNSGSDAST--NPDESAGPNLTVISKRTDSNA---FL 414
DB 1633 DINELEIA-----LDHANANADAOKNLKRQYEQVRELQYEEBQRNGADTREQFFN 1685
OY 415 AKREVEALLSSIDEI--SKAIGKIKKNDGTLD-----NEANRNSLAGA----- 457
DB 1686 AKRATILLQSKKEELVLANEAERAKQAEYEAADQANEAQVSSLTSKARKLEGE 1745
OY 458 -----YEISKITITQ-----KLVSUNSEELKKRIKAEKDCSKFTTKIKDSHAELGIQ 504
DB 1746 IQATHADDETITNMYKKAERSKKAIDATRLAELEHQEQEHSHQVDRLLKGLGQQLKEI 1805
OY 505 SYODDNKKKAILKTHGKDKGAKLELEFSLSLSKAAQALITNSYKEL 554
DB 1806 QVRLEDEAALAK--GGKKVIAR--LEQVRRELESELDGEORRPODANKNL 1852

RESULT 13
SCA4_RICPR STANDARD; PRT; 1022 AA.
ID 092D49: 092D48; 09AJ36;
AC 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE ANTIMICIN HEAT-STABLE 120 KDA PROTEIN (P5120) (120 KDA ANTIGEN)
DE (PROTEIN PS 120).
GN SCA4 OR RP498/RP499.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
RN [2]
RP SEQUENCE OF 11-1016 FROM N.A.
RA Sekeyova Z., Roux V., Raoult D.;
RT "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the
RT 'gene D' coding for an intracytoplasmic protein."
RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 234.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AJ235272; CA114951.1; ALT_FRAME.
DR EMBL: AJ235272; CA114950.1; ALT_FRAME.
DR EMBL: AF200340; AAK31305.1;
KW Antigen; Complete proteome.
FT CONFLICT 11 15
FT CONFLICT 365 365 EFDPL -> Y RPYLV (IN REF. 2).
FT CONFLICT 413 413 MISSING (IN REF. 2).
FT CONFLICT 957 957 G -> R (IN REF. 2).
FT SEQUENCE 1022 AA; 114410 MW; 03230E3A663A9622 CRC64;

```

Query Match 6.9%; Score 189; DB 1; Length 1022;
 Best Local Similarity 23.5%; Pred. No. 0.25;
 Matches 158; Conservative 89; Mismatches 204; Indels 220; Gaps 37;

```

OY 5 NSGKGNNTSANSNAD--SYKGNLFEISKRTSDSNVLAKEVEALLSIDELIA-----58
DB 322 NKGOSVALSQTIAEDLTTHVQGPS-HETNRPILIPN-----QELSSIEQHTSQOVP 371
OY 59 -----KAIGKKIHONGNGDTEYNNHNGSL-----LAGAVAISTLIKOKLDG-----KNEG 103
DB 372 PTTTNNKSKQPKISQIHQOPOAOSSGIPNPVLNANALSTSMODLNNINSYLRKND 431
OY 104 LKEKIDAAKKCSEPTNKLK-----EKHTDL-----GREGVTDADA-----139
DB 432 INKOSDLIEKAIAIILNNKSPFAEKQYNIIDLAKNIFSNKDIIADAKVAVVLTLETIQ 491
OY 140 -----KEALIKTNGTKTGAELGKLFESVEVLSKAAREMLANSVKELTSPVAES 190
DB 492 NDONTLIRKSKRIEDTVAITLNSENI-ELKOKQOILEKVVD--IGLSIKDISRYVAVD 548
OY 191 PAMGNSGKGDSTNPADDESAGKPNLT-----EIS-----KRTTDSNAFLAVKEV 238
DB 549 SIMDT-----VKSININEDKEKIFITVFDQINSYFSNAKOKLDS-----ILKKA 597
OY 239 ETLVLISIDELAKKAIGOKIDN-----NNGI-----AALNNGSLAGAY 278
DB 598 ETVOVLSPEQ--QOLMNQNDNITTEHTKRDTEKVNILLEPLSNALAKTTNIQVMTSNV 655
OY 279 AISTLITELSKL-----KNEEL-----KTEIAKAKKCEEFNKLKSGHAGDQKD 326
DB 656 LDPVOIEKMSKRLIOVTVTAESALVEBKDTETIYKG-----IGKTI 698
OY 327 ATDD-----HAKAAILKTHATTDKGAKEFKDFESVEGLKKAQY--ALTNSVKELGHR 378
DB 699 VTHSDTSLPLHDKVYIMGSA-----KGIYESKNLDLRELIAGLVDTIYE-----745
OY 379 NNSGDSASTNPDESAGKPNLTV--ISKKITDSNAFLAVKEVEALLSIDELS-RAIGK 435
DB 746 -----AKGNNAVVAHAISSMIANSN--INQSEKALKRSODVYSEKVLDK 787
OY 436 KIKNDGTLDNE--ANRNSLIAG-----AYEISKLITQKLSVLN-----SEC 475
DB 788 ETQN--LDRELKAQININSKLDHDIYNTQOVANALKVVIT--TVLDNSGQSGVSEE 841
OY 476 LKKKIKE-AKDCSQKFTTKLKDASHAELGQSVODDAAKKAILKT-HGTRDKGAKELEEL 533
DB 842 ARKVVSSLLNDISKRTIEKINLRLMALS-----QDGN-----LKTFEKKDEATKRYVDLY 892
OY 534 KSELEISKAQ 544
DB 893 KAFDNKSSTEE 903

```

```

RA Bilbe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA Cerletti N., Sorg C., Odink K., Tarcsey L., Wiesendanger W.,
RA de Wolf-Peters C., Shipman R.,
RA "Restin: a novel intermediate filament-associated protein highly
RA expressed in the Reed-Sternberg cells of Hodgkin's disease.";
RA EMBO J. 11:2103-2113(1992).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE-92405160; PubMed-1356075;
RA Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
RA "Clp-170 links endocytic vesicles to microtubules.";
RA Cell 70:887-900(1992).
CC CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULS.
CC CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC CC CYTOSKELETON.
CC CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERBERG CELLS
CC CC OF HODGKIN'S DISEASE.
CC CC -1- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: X64838; CAA6050.1; -.
DR EMBL: M97501; AAA5693.1; -.
DR PIR: S22695; S22695.
DR MIM: 179838; -.
DR InterPro: IPR000938; CAP-GLY.
DR InterPro: IPR001878; znf.CCHC.
DR Pfam: PF01302; CAP_GLY; 2.
DR SMART: SM00343; znf.C2HC; 1.
DR PROSITE: PS00845; CAP_GLY_1; 2.
DR KEGG: Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT DOMAIN 60 125 CAP-GLY 1.
FT DOMAIN 143 204 SER-RICH.
FT DOMAIN 214 279 CAP-GLY 2.
FT DOMAIN 304 331 SER-RICH.
FT DOMAIN 350 342 COILED COIL (POTENTIAL).
FT DOMAIN 1408 1421 CCHC-BOX.
FT VARSPIC 457 491 MISSING (IN SHORT ISOFORM).
FT CONFLICT 1069 1069 D -> E (IN REF. 2).
SQ :SEQUENCE 1427 AA; 160989 MW; 0A4F16DD94254E8 CRC64;

```

Query Match 6.9%; Score 189; DB 1; Length 1427;
 Best Local Similarity 21.6%; Pred. No. 0.36;
 Matches 135; Conservative 100; Mismatches 249; Indels 140; Gaps 26;

```

OY 28 EISKKITDSNAVLAVKEVEALLSIDELIAKKAIGKTIHONGND-----TEYNNGS 80
DB 523 ESNKPGADVDMSTSLQLEISSIQEKE-----VTRDHOREITSLKEHNGARETHQKE 576
OY 81 LLAGAVAISTLIKOKLDGKLNKGAKEKIDAKKCEFTNKLKLEK-HTDQKSGVTDADA 139
DB 577 IKALVYATEKLKSE-----NESIKSKLEHANKENSVDVIALMSKLETAISHQOAMEEL 630
OY 140 KEAAILKNTNGTKTGAELGKLFESVEVLSKAAREMLANSVKELTSPVAESPAWGSN---196
DB 631 KVSFSGGLGTEPAELFAELKQIETKMRDYOHEIENLQNDQSEARAHAKEALRAKLAKM 690
OY 197 --SGKGDASTNPADDESAGKPNLTETSKITTSNAFLAVKEVEVTLVLSIDELAK-----250
DB 691 VIREKENSLEAIRSKLDKADQHLVEMEDPLNKQEAIEIVKLEVLQAKCNQOTKVIDN 750
OY 251 -----RAIGOKI-----DNNNGLAALNNONGSLAGAVAI STL-----I 284

```

```

Db 751 FTSQKATEKLLDLDLRKASSEKSEKMLKQOUEAAEKOIHLEIEKNASSKASSI 810
Qy 285 TEKLS-----KLKLELKEITKAIKAKKSEFTNKLSGSHADIGKOATDDHA-KKAILKT 339
Db 811 TRELOGRELKTNTQENLSESVQVETLEKEIQLIKKEKAESESVQSRMOETVNL 870
Qy 340 HA-----TTDKGAKERKDLFESEVGLLKAQVALTNSVKELGHRNNS 381
Db 871 HCKEOPNMLSSDLKRLRENLADEAK-FREKDEREQLITAKE-KLENDIAEI---MKM 925
Qy 382 GGDSSA-----TTPDESAGKPNLTVISKTTIDSNAPLAVKEVALSSIDELSKAIGKKI 437
Db 926 SGDNSQSLTKMDELKRLERDVEELQKLTJKN-----ENASFLQKSIDMT-----V 973
Qy 438 KNDGTLDNEANRNSLIGAVEISKLTQKLSVLNSEELKKI-----KEAKDSQKF 490
Db 974 KAEQSQEAKKHE-----EKKELERKLS-----DEKKEHSHOCCOLKARYERA 1021
Qy 491 TTKLDSHAELGIGSVODNNAKALIKTHGTQKQAK-----ELELEKSLSELSKA 542
Db 1022 TSETTKHEE-LQNLQ-----KTLIDTE-DKLKGRENSGLLOLELELRKQADK-AKA 1073
Qy 543 AQA-----LTNSVKELTNPVVA 560
Db 1074 AQTADDAQIMEQMKREKTEITLAS 1097

```

```

RESULT 15
Z1P1_YEAST STANDARD: PRT: 875 AA.
AC P31111:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SYNAPTONEMAL COMPLEX PROTEIN Z1P1.
GN Z1P1 OR YDR285W OR D9819.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

```

```

RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BR1824-3B;
RX MEDLINE=93161412; PubMed=7916652;
RA Sym M., Engelbrecht J.A., Roeder G.S.;
RA "Z1P1 is a synaptonemal complex protein required for meiotic
chromosome synapsis."
Cell 72:365-378(1993).

```

```

RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Johnson M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Pavello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifken L., Riles L., Taich A., Trevisan E., Vignati D.,
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

```

```

CC -!- FUNCTION: REQUIRED FOR MEIOTIC CHROMOSOME SYNAPSIS AND CELL CYCLE
PROGRESSION. MAY ACT AS A MOLECULAR ZIPPER TO BRING HOMOLOGOUS
CHROMOSOMES IN CLOSE APPPOSITION. Z1P1 MAY ENCODE THE TRANSVERSE
FILAMENTS OF THE SYNAPTONEMAL COMPLEX.

```

```

CC -!- SUBCELLULAR LOCATION: SYNAPSED MEIOTIC CHROMOSOMES.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: L06487; AAA35239.1; -;
DR EMBL: U51031; AAB64474.1; -;
DR PIR: A45173; A45173.
DR PIR: S30868; S30868.
DR SGD: S0002693; Z1P1.
KW Nuclear protein; Meiosis; Coiled coil.
FT DOMAIN 177 333 COILED COIL (POTENTIAL).
FT DOMAIN 397 438 COILED COIL (POTENTIAL).
FT DOMAIN 456 752 COILED COIL (POTENTIAL).
FT CONFLICT 55 55 T -> A (IN REF. 1).
SQ SEQUENCE 875 AA; 100035 MW; 674F12625CD9DDFD CRC64;

```

Query Match 6.8%; Score 187; DB 1; Length 875;

Best Local Similarity 20.18; Pred. No. 0.25;

Matches 115; Conservative 117; Mismatches 197; Indels 142; Gaps 25;

```

Qy 20 SVKGPNTL--EISKRTDSNAVLAV--KEVEALLSIDETIAAKAIGKRIHQNGGLDTEYN 76
Db 274 SIEMLNTINDLGKRRKADAELEKKGKEIEYLRKRELDGSGQSEKTI-KNSSLIQEMG 332
Qy 77 HNSLLGATAISTLIKQKIDGLKNEGKTKI-----DAKKSETFTNKLKEKHTDL 129
Db 333 KNR-----EEMKRSIENFSEDKAHLLQFNKFEERVHDLFEKKL-QKHFDV 378
Qy 130 GKEGVTDADAKKALIKNGTKTKGAELGKLFESVEVLSKAKEMLANSVKELTSPVAV 189
Db 379 AKDTLVNGLNTYVELSSNTETMLKQOYEDIKENLEOKMSSKDEAKKTINEIS--VTK 436
Qy 190 SPANGSGSGKGGDSASTNPDESAGKPNLTVISKTTIDSNAPLAVKEVELVLSIDELA 249
Db 437 GLIMG-----VOEELITSG-----NIOTALVS----- 459
Qy 250 KKAIGKIDNNGLALNNNGSLAGAVASITLIEKSLKLTN--ELKTEIARAKKS 308
Db 460 -----EKNNTROELDDA-----SQTAKNYSLENLVAKVKAETVOSNE-Y 499
Qy 309 EEFNTKLSGSHADLGQDATDHAKAAILKTHATTDKAKE--FKDLFESVEGLKKAQV 366
Db 500 EERIKHLESERSTLSSQ-----KNQIISLGRKQAYELDVAKKLE--KNIEI 545
Qy 367 A-LTNSVKELGHHNNNGSGDSASTNPDESAGKPNLTVISK-----KTTDSNAPLAVKEVE 420
Db 546 SQISGKEQSLTEKENNLSNELKKRVQDLERKLNLTITTSYENKISSONEI-----YK 599
Qy 421 ALLSIDELSKAIGKKI-----KNDGTLDNEA--NRNSLIGAVEISKLTQKLSVLN 472
Db 600 ALVSENDTLKRIQQLVEIKENEDKHDTKLEAFQKNNQDLQKLVNEVQVLAHEL----- 655
Qy 473 SEFLKKIKKAEKDCSQKFTTKLKDASHAELGI--QSV-----QDNNAKALIKTHGTRD 523
Db 656 --ELEEDONHKLKCLEKKEKTEGVEESLSDVTKLQOYIVLSKODITABKLEQON---- 709
Qy 524 KGAKLEELFKSLDESLSKAAQVALTNSVKE 554
Db 710 --LESLEEVTKNLQKVVQSKRELQKIKEL 738

```

Search completed: March 18, 2002, 10:12:03
Job time: 981 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:55 ; Search time 124.19 seconds

(without alignments)
659.574 Million cell updates/sec

Title: US-09-596-746a-52

Perfect score: 2750
Sequence: 1 MACNNSGKNGNTSANSADSE.....KAAQALNTSVKELTNPVVA 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by the result to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	926	33.7	192	2	Q9S3P3
2	921	33.5	192	2	Q9R7B1
3	910	33.1	191	2	P70818
4	908	33.0	193	2	Q9R853
5	893	32.5	200	2	Q9R0R9
6	886	32.2	194	2	Q9S3P4
7	885	32.2	182	2	Q9R7B2
8	884	32.1	203	2	P96507
9	881	32.0	212	2	Q07336
10	880	32.0	207	2	Q9K1M4
11	878	31.9	207	2	Q45187
12	878	31.9	207	2	Q9REPF8
13	878	31.9	207	2	Q9REPF7
14	877	31.9	212	2	Q44669
15	876	31.9	203	2	Q50623
16	874	31.8	200	2	Q08139
17	853	31.0	177	2	Q44999
18	846	30.8	179	2	Q44987
19	844	30.7	179	2	Q44988

20	843.5	30.7	178	2	Q9REH7
21	834.5	30.3	180	2	Q44991
22	832	30.3	179	2	Q44986
23	829	30.1	175	2	Q9R7B4
24	805	29.3	189	2	P94224
25	802	29.2	163	2	Q9R7A9
26	798	29.0	207	2	Q45177
27	794	28.9	200	2	Q44996
28	786	28.6	207	2	Q49581
29	773	28.1	159	2	Q9R7B0
30	759.5	27.6	199	2	Q44998
31	757.5	27.5	211	2	Q44720
32	756.5	27.5	193	2	P94237
33	755.5	27.5	199	2	Q08233
34	745.5	27.1	204	2	P96509
35	705.5	25.7	206	2	P96512
36	700	25.5	210	2	Q44719
37	698	25.4	192	2	Q9S3P2
38	696.5	25.3	193	2	P94234
39	691	25.1	193	2	Q31115
40	687.5	25.0	204	2	Q44989
41	683.5	24.9	211	2	Q49576
42	678.5	24.7	194	2	Q31122
43	678.5	24.7	211	2	Q44977
44	677	24.6	212	2	Q9K1M5
45	676.5	24.6	209	2	Q44671

ALIGNMENTS

RESULT	ID	Q9S3P3	PRELIMINARY:	PRT:	192 AA.
AC	Q9S3P3				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			
DE	OUTER SURFACE PROTEIN C (FRAGMENT).				
GN	OSPC.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.				
OX	NCBI_TaxID=139;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IP2;				
RX	MEDLINE=96296448; PubMed=8709845;				
RA	Lively I., Gibbs C.P., Schuster R., Dornier F.;				
RT	"Evidence for lateral transfer and recombination in OspC variation in				
RL	Lyme disease Borrelia."				
DR	EMBL: L42887; AAB36995.1; "				
DR	InterPro: IPR001800; Lipoprotein_6.				
DR	Pfam: PF01441; Lipoprotein_6; 1.				
FT	Prodom: PD001149; Lipoprotein_6; 1.				
FT	NON_TER	1			
FT	NON_TER	192			
SQ	SEQUENCE	192 AA; 20287 MW; 118467AC84CTE3D CRC64;			
Query Match					
Best Local Similarity 99.5%; Score 926; DB 2; Length 192;					
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
OY	3	CNNSGKDGNTSANSADSEVKGPNLTETSKITPDSNVLLAVKVEALLSIDIAKAIG 62			Q9REH7 borrelia bu
DB	1	CNNSGKDGNTSANSADSEVKGPNLTETSKITPDSNVLLAVKVEALLSIDIAKAIG 60			Q44991 borrelia af
OY	63	KTIHNNGLDTEVYNNHNSLAGAVALSTLIKRLDIKNEGLEKIDAAKCKSEFTTNKL 122			Q44986 borrelia af
DB	61	KTIHNNGLDTEVYNNHNSLAGAVALSTLIKRLDIKNEGLEKIDAAKCKSEFTTNKL 120			Q9R7B4 borrelia bu
OY	123	KEHTDGLGEGYVDADAKELIKTNGTKTGAEELGKLPESVVLAKAKEMLANVSKEL 182			Q9R7A9 borrelia bu

```

Db 121 KEKHTDLGKRGVTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAKKEMLANSVKEL 180
QY 183 TSPVVAESP 191
Db 181 TSPVVAESP 189

RESULT 2
Q9R7B1
ID 09R7B1 PRELIMINARY: PRT: 192 AA.
AC 09R7B1
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
NCBI_TaxID=139;
[1]
RC SEQUENCE FROM N.A.
RA MEDLINE=97478003; PubMed=9336916;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
  U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U91798; AAB81895.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 20297 MW; 6770502A20AA764 CRC64;

Query Match 33.5%; Score 921; DB 2; Length 192;
Best Local Similarity 98.4%; Pred. No. 6.6e-35;
Matches 187; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSEYKGNLFEISKRTDSNAVLAVKEVALLSSIDEIAKA 60
   ::::|||||
3 ISCNNSGKDGNTSANSADSEYKGNLFEISKRTDSNAVLAVKEVALLSSIDEIAKA 62
QY 61 IGGKIHQNGLDTEYNHNGSLAGAVAI STLIRKOKLDGKNEGLKEKIDAAKCSFTFTN 120
   :|||||
Db 63 IGGKIHQNGLDTEYNHNGSLAGAVAI STLIRKOKLDGKNEGLKEKIDAAKCSFTFTN 122
QY 121 KLAKEKHTDLGKRGVTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAKKEMLANSVK 180
   :|||||
Db 123 KLAKEKHTDLGKRGVTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAKKEMLANSVK 182
QY 181 ELTSPVVAES 190
   :|||||
Db 183 ELTSPVVAES 192

RESULT 3
P70818
ID P70818 PRELIMINARY: PRT: 191 AA.
AC P70818
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DE OUTER SURFACE PROTEIN (FRAGMENT).
OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

```

```

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2-1498 CA4.
RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U81131; AAB06569.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 191
SQ SEQUENCE 191 AA; 20126 MW; D2B9B1C82B4DC3C0 CRC64;

Query Match 33.1%; Score 910; DB 2; Length 191;
Best Local Similarity 98.9%; Pred. No. 2.1e-34;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 NNSGKDGNTSANSADSEYKGNLFEISKRTDSNAVLAVKEVALLSSIDEIAKAIGK 63
   :|||||
Db 1 NNSGKDGNTSANSADSEYKGNLFEISKRTDSNAVLAVKEVALLSSIDEIAKAIGK 60
QY 64 KIHQNGLDTEYNHNGSLAGAVAI STLIRKOKLDGKNEGLKEKIDAAKCSFTFTN 123
   :|||||
Db 61 KIHQNGLDTEYNHNGSLAGAVAI STLIRKOKLDGKNEGLKEKIDAAKCSFTFTN 120
QY 124 EKHTDLGKRGVTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAKKEMLANSVKELT 183
   :|||||
Db 121 EKHTDLGKRGVTDADAKKALIKTNGTKTGAEELGKLFESVEVLSKAKKEMLANSVKELT 180
QY 184 SPVVAESP 191
   :|||||
Db 181 SPVVAESP 188

RESULT 4
Q9R53
ID Q9R53 PRELIMINARY: PRT: 193 AA.
AC Q9R53
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSCP.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OC1;
RX MEDLINE=99091544; PubMed=9872945;
RA Wang I.N., Dykhuizen D.E., Olu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT "Genetic diversity of oSCP in a local population of Borrelia
  burgdorferi sensu stricto.";
RL Genetics 151:15-30(1999).
DR EMBL: AF029860; AAB86543.1;
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 20502 MW; 5EFD85AF8986D1E CRC64;

Query Match 33.0%; Score 908; DB 2; Length 193;
Best Local Similarity 98.4%; Pred. No. 2.6e-34;
Matches 184; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MACNNSGKDGNTSANSADSEYKGNLFEISKRTDSNAVLAVKEVALLSSIDEIAKA 60
   :|||||
Db 7 ISCNNSGKDGNTSANSADSEYKGNLFEISKRTDSNAVLAVKEVALLSSIDEIAKA 66
QY 61 IGGKIHQNGLDTEYNHNGSLAGAVAI STLIRKOKLDGKNEGLKEKIDAAKCSFTFTN 120

```



```

|||||
Db 67 ICKKHONNGLDTENNHNSLLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 126
OY 121 KKEKHNDLGKGVTDADAKKAEILKTNGTKTKGAEELGKLFESEVLSAAKEMLANSVK 180
Db 127 KKEKHNDLGKGVTDADAKKAEILKTNGTKTKGAEELGKLFESEVLSAAKEMLANSVK 186
OY 181 ELTSPV 187
Db 187 ELTSPV 193

RESULT 5
O9ROR9 PRELIMINARY: PRT: 200 AA.
ID O9ROR9:
AC O9ROR9: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OSPC.
Borrelia burgdorferi (Lyme disease spirochete).
Plasmid cp26.
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2E7;
RX Homeister E.K., Glass G.E., Childs J.E., Persing D.H.:
  "Population dynamics of a naturally occurring heterogeneous mixture of
  Borrelia burgdorferi clones."
  Infect. Immun. 67:5709-5716(1999).
RL EMBL: AF074464; AAD3911.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
KM Plasmid.
FT NON_TER 200
SQ SEQUENCE 200 AA; 21307 MW; A22F158758BBEB6B CRC64;

Query Match 32.5%; Score 893; DB 2; Length 200;
Best Local Similarity 98.4%; Pred. No. 1.3e-33;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 MACNNSGKGGNTSANSADSVKGPNTLEISKRTTDSNAVLAVKEVALLSSTIDEIAAKA 60
Db 17 ICKKHONNGLDTENNHNSLLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 76
OY 61 ICKKHONNGLDTENNHNSLLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 120
Db 77 ICKKHONNGLDTENNHNSLLAGAVAI STLKOKLDGLKNGELKIDAAKCCSEFTFN 136
OY 121 KKEKHNDLGKGVTDADAKKAEILKTNGTKTKGAEELGKLFESEVLSAAKEMLANSVK 180
Db 137 KKEKHNDLGKGVTDADAKKAEILKTNGTKTKGAEELGKLFESEVLSAAKEMLANSVK 196
OY 181 ELTSPV 184
Db 197 ELTSPV 200

RESULT 6
O9S3P4 PRELIMINARY: PRT: 194 AA.
ID O9S3P4:
AC O9S3P4: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
Borrelia burgdorferi (Lyme disease spirochete).

```

```

OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JSB;
RX MEDLINE=96296448; PubMed=8709845;
RA Livey I., Gibbs C.P., Schuster R., Dörner F.:
  "Evidence for lateral transfer and recombination in OspC variation in
  Lyme disease Borrelia."
  Mol. Microbiol. 18:257-269(1995).
RL EMBL: L42883; AAB36991.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 194
SQ SEQUENCE 194 AA; 20446 MW; CEEDC9FA5DF0D68F CRC64;

Query Match 32.2%; Score 886; DB 2; Length 194;
Best Local Similarity 90.7%; Pred. No. 2.5e-33;
Matches 186; Conservative 2; Mismatches 3; Indels 14; Gaps 2;
OY 195 SNSGKGSASSTNPADSAKGPVLTETSKRTTDSNAFVLAVKEVTLVSTIDEIAKKAIG 254
Db 2 NNSGKGSASSTNPADSAKGPVLTETSKRTTDSNAFVLAVKEVTLVSTIDEIAKKAIG 61
OY 255 OKIDNNNGLAALNONGSLAGAVAI STLTKSLKLNLEELKTEIAAKKCCSEFTNK 314
Db 62 OKIDNNNGLAALNONGSLAGAVAI STLTKSLKLNLEELKTEIAAKKCCSEFTNK 121
OY 315 LKSGHADLKODATDDHAKAAILKTHTATDKAKERKDLFESEGLKKAQVALTNSVKE 374
Db 122 LKSGHADLKODATDDHAKAAILKTHTATDKAKERKDLFESEGLKKAQVALTNSVKE 181
OY 375 LGRHNSGGSASTNP--DESAKGP 397
Db 182 L-----TSPVAESPAPK 194

RESULT 7
O9R7B2 PRELIMINARY: PRT: 182 AA.
ID O9R7B2:
AC O9R7B2: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.:
  "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
  U.S.A.'?"
  Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TETS;
RA Marti-Ras N., Postic D., Foretz M., Baranton G.:
  Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: U91797; AAB81894.1; -.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 182
SQ SEQUENCE 182 AA; 19202 MW; 422146F99A57BF2 CRC64;

```

Query Match 32.2%; Score 885; DB 2; Length 182;
Best Local Similarity 99.5%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 7 GKQNTSANSADSVKPNLTETISKTTDSNAVLAVKEVEALLSIDETAKAIGKIH 66
DB 1 GKQNTSANSADSVKPNLTETISKTTDSNAVLAVKEVEALLSIDETAKAIGKIH 60
QY 67 QNNGLDTEYHNHNSLAGAVAIATLTKOKLDGKNGEKIDAAKCKSETFTNKLEKH 126
DB 61 QNNGLDTEYHNHNSLAGAVAIATLTKOKLDGKNGEKIDAAKCKSETFTNKLEKH 120
QY 127 TDGKGVYDADAKKAIKLTNGTKGAEELKLFESVEVLSKAKEMLANSVKELTSPV 186
DB 121 TDGKGVYDADAKKAIKLTNGTKGAEELKLFESVEVLSKAKEMLANSVKELTSPV 180
QY 187 VA 188
DB 181 VA 182
```

RESULT 8

```
ID P96507 PRELIMINARY; PRT: 203 AA.
AC P96507;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSpC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BF03;
RA Masuzawa T., Komikado T., Fukui T., Yanagihara Y.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB000985; BAA24125.1; -
DR EMBL: AF001800; Lipoprotein_6.
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 203
SQ SEQUENCE 203 AA; 21549 MW; 94A885FBBACEF1C4 CRC64;
```

Query Match 32.1%; Score 884; DB 2; Length 203;
Best Local Similarity 99.4%; Pred. No. 3.3e-33;
Matches 180; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 195 SNSGKGSASTNPADSAGPNLTETISKTTDSNAVLAVKEVEALLSIDELAKKAIG 254
DB 14 SNSGKGSASTNPADSAGPNLTETISKTTDSNAVLAVKEVEALLSIDELAKKAIG 73
QY 255 OKIDNNNGSLALNNONGSLAGAVAIATLTKETLSKLNLEELKTELAKKCKSEFTNK 314
DB 74 OKIDNNNGSLALNNONGSLAGAVAIATLTKETLSKLNLEELKTELAKKCKSEFTNK 133
QY 315 LKSGHADLQKQATDTHAKAAIILKTHATTDKGAKEFKDLFESVEGLLKAQVALTNSVKE 374
DB 134 LKSGHADLQKQATDTHAKAAIILKTHATTDKGAKEFKDLFESVEGLLKAQVALTNSVKE 193
QY 375 L 375
DB 194 L 194
```

RESULT 9

```
ID Q07336 PRELIMINARY; PRT: 207 AA.
AC Q07336;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C.
GN OSpC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PB1;
RA MEDLINE-93268136; PubMed-8098841;
RA Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA Will G., Wilske B.;
RT "Genetic heterogeneity of the genes coding for the outer surface
RT protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL Med. Microbiol. Immunol. 182:37-50(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PB1;
RA MEDLINE-93293332; PubMed-8478108;
RA Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA Schwab E., Wanner G.;
RT "Immunological and molecular polymorphisms of OSpC, an immunodominant
RT major outer surface protein of Borrelia burgdorferi.";
RL Infect. Immun. 61:2182-2191(1993).
DR EMBL: X69595; CAA49305.1; -
DR InterPro: IPR001800; Lipoprotein_6.
DR Pfam: PF01441; Lipoprotein_6; 1.
DR ProDom: PD001149; Lipoprotein_6; 1.
SQ SEQUENCE 207 AA; 22320 MW; 4A525ABDAE63B5C8 CRC64;
```

Query Match 32.0%; Score 881; DB 2; Length 207;
Best Local Similarity 99.5%; Pred. No. 4.6e-33;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 379 SNSGGSASTNPDESAGPNLTETISKTTDSNAVLAVKEVEALLSIDELSAIGKIK 438
DB 20 SNSGGSASTNPDESAGPNLTETISKTTDSNAVLAVKEVEALLSIDELSAIGKIK 79
QY 439 NDGTLNDEANRNESLJAGAVEISKLITTKLSVNSEELKKIKKEAKDCSQKFTTKLDKSH 498
DB 80 NDGTLNDEANRNESLJAGAVEISKLITTKLSVNSEELKKIKKEAKDCSQKFTTKLDKSH 139
QY 499 AELGISVDDNNKKAIIKTHGKDKGAKELDFRSLSSLSKAAQALTNVSKELTNPV 558
DB 140 AELGISVDDNNKKAIIKTHGKDKGAKELDFRSLSSLSKAAQALTNVSKELTNPV 199
QY 559 VA 560
DB 200 VA 201
```

RESULT 10

```
ID Q9KIM4 PRELIMINARY; PRT: 212 AA.
AC Q9KIM4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C.
GN OSpC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN-PCAU:
 RA Dykhuizen D.E., Gutman D.S., Luft B.J.:
 RT "Antigenic variation and integrative recombination in the ospC gene of
 RT Borrelia burgdorferi.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF230185; AAF75626.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 SQ SEQUENCE 212 AA; 22495 MW; C5IDAF803BF2E7D4 CRC64;

Query Match 32.0%; Score 880; DB 2; Length 212;
 Best Local Similarity 90.2%; Pred. No. 5.2e-33;
 Matches 185; Conservative 2; Mismatches 4; Indels 14; Gaps 2;

QY 195 NSNGGGSASTNPDESAGKPNLTETSKRTDSDNAFLAVKEVELVLSIDELAKKAIG 254
 :|||||
 DB 20 NSNGGGSASTNPDESAGKPNLTETSKRTDSDNAFLAVKEVELVLSIDELAKKAIG 79

QY 255 QKIDNNNGSLAALNONGSLAGAYASTLITETKSLKMLELKTETIAKAKCSEFTTK 314
 :|||||
 DB 80 QKIDNNNGSLAALNONGSLAGAYASTLITETKSLKMLELKTETIAKAKCSEFTTK 139

QY 315 LKSGHADLCKODATDPAKATILKTHATDKGAKKEFKDLFESVEGLKAAVALTNSVKE 374
 :|||||
 DB 140 LKSGHADLCKODATDPAKATILKTHATDKGAKKEFKDLFESVEGLKAAVALTNSVKE 199

QY 375 LGRNNSGGSASTNP--DESAKGP 397
 :|||
 DB 200 L-----TSPVVAESPCKP 212

RESULT 11
 Q45187 PRELIMINARY; PRT; 207 AA.

AC Q45187;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN C PRECURSOR.
 GN OSPC.
 OS Borrelia garinii.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID:29519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PTROB;
 RX MEDLINE-95213332; PubMed-7699024;
 RA Wilske B., Jauris-Helpe S., Lobentanz R., Pradel I.,
 RA Preac-Mursic V., Roseler D., Soutschek E., Johnson R.C.;
 RA Phenotypic analysis of outer surface protein C (OspC) of Borrelia
 RT burgdorferi sensu lato by monoclonal antibodies: relationship to
 RT genospecies and OspA serotype.";
 RL J. Clin. Microbiol. 33:103-109(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PTROB;
 RX MEDLINE-95395018; PubMed-7665660;
 RA Jauris-Helpe S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,
 RA Soutschek E., Will G., Wilske B.;
 RA "Molecular analysis of genes encoding outer surface protein C (OspC)
 RT of Borrelia burgdorferi sensu lato: relationship to ospA genotype and
 RT evidence of lateral gene exchange of ospC.";
 RL J. Clin. Microbiol. 33:1860-1866(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCSC, PBAETI, PFIN, PFUK, PMUE, PSB;
 RA Marconi R.T., Hohenberger S., Jauris-Helpe S., Schulte-Spechtel U.,
 RA Lavole C.P., Roessler D., Wilske B.;
 RA "Genetic analysis of B.garinii OspA-serotype 4 strains associated with
 RT neuroborreliosis: evidence for extensive genetic homogeneity.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: X83554; CAA58544.1; -
 DR EMBL: AJ236908; CAB46238.1; -
 DR EMBL: AJ132793; CAB46231.1; -
 DR EMBL: AJ132796; CAB46234.1; -
 DR EMBL: AJ132797; CAB46235.1; -
 DR EMBL: AJ132798; CAB46236.1; -
 DR EMBL: AJ236907; CAB46237.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 KW Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 207
 SQ SEQUENCE 207 AA; 22321 MW; 20889AA76E63A49D CRC64;

Query Match 31.9%; Score 878; DB 2; Length 207;
 Best Local Similarity 98.9%; Pred. No. 6.2e-33;
 Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NSNGGGSASTNPDESAGKPNLTETSKRTDSDNAFLAVKEVELVLSIDELSKAIGKK 438
 :|||||
 DB 20 NSNGGGSASTNPDESAGKPNLTETSKRTDSDNAFLAVKEVELVLSIDELSKAIGKK 79

QY 439 NDSGLTNEARNRNSLJAGAYEISKLITOKLSVNSEELKKIKKAKDCSOKFTTKLKD 498
 :|||||
 DB 80 NDSGLTNEARNRNSLJAGAYEISKLITOKLSVNSEELKKIKKAKDCSOKFTTKLKD 139

QY 499 AELGISOVDNNAKKAILKTHGTRDKGAKLEELFSLSLSKRAQAALNSVKELTNPV 558
 :|||||
 DB 140 AELGISOVDNNAKKAILKTHGTRDKGAKLEELFSLSLSKRAQAALNSVKELTNPV 199

QY 559 VA 560
 :||
 DB 200 VA 201

RESULT 12
 Q9REF8 PRELIMINARY; PRT; 207 AA.

AC Q9REF8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE OUTER SURFACE PROTEIN (FRAGMENT).
 GN OSPC.
 OS Borrelia garinii.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID:29519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PWA;
 RX MEDLINE-20033636; PubMed-10565915;
 RA Marconi R.T., Hohenberger S., Jauris-Helpe S., Schulte-Spechtel U.,
 RA Lavole C.P., Roessler D., Wilske B.;
 RA "Genetic analysis of Borrelia garinii OspA serotype 4 strains
 RT associated with neuroborreliosis evidence for extensive genetic
 RT homogeneity.";
 RL J. Clin. Microbiol. 37:3965-3970(1999).
 DR EMBL: AJ132794; CAB46232.1; -
 DR InterPro: IPR001800; Lipoprotein_6.
 DR Pfam: PF01441; Lipoprotein_6; 1.
 DR ProDom: PD001149; Lipoprotein_6; 1.
 FT NON-TER 207 207
 SQ SEQUENCE 207 AA; 22366 MW; 30CFEAA6D913A49D CRC64;

Query Match 31.9%; Score 878; DB 2; Length 207;
 Best Local Similarity 98.9%; Pred. No. 6.2e-33;
 Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 379 NSNGGGSASTNPDESAGKPNLTETSKRTDSDNAFLAVKEVELVLSIDELSKAIGKK 438
 :|||||
 DB 20 NSNGGGSASTNPDESAGKPNLTETSKRTDSDNAFLAVKEVELVLSIDELSKAIGKK 79

```

DB 20 NNSGSDASTNPDESAGKPNLTVISKITDTSNAFLAVKEVEALLSIDELSKAIGKKIK 79
QY 439 NDGTLDNNEANRNESILAGAVEISKLTOKLSVLSNSELKKIKKEAKDCSOKFTTKLDSH 498
DB 80 NDGTLDNNEANRNESILAGAVEISKLTOKLSVLSNSELKKIKKEAKDCSOKFTTKLDSH 139
QY 499 AELGISVODDANKAKILKTHGTGDKGAKELFEKLSLESLSKAQAALNTSVKELTNPV 558
DB 140 AELGISVODDANKAKILKTHGTGDKGAKELFEKLSLESLSKAQAALNTSVKELTNPV 199
QY 559 VA 560
DB 200 VA 201

RESULT 13
QREF7 09REF7 PRELIMINARY: PRT: 207 AA.
01-MAY-2000 (TREMBLrel. 13, Created)
01-JUN-2001 (TREMBLrel. 17, Last sequence update)
OUTER SURFACE PROTEIN.
OSPC.
Borrelia garinii.
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=29519;
SEQUENCE FROM N.A.
STRAIN-PRO;
MEDLINE=20033636; PubMed=10565915;
Marconi R.T., Hoenesberger S., Jauris-Helpe S., Schulte-Spechtel U.,
LaVoie C.P., Roessler D., Wilske B.;
"Genetic analysis of Borrelia garinii Ospa serotype 4 strains
associated with neuroborreliosis evidence for extensive genetic
homogeneity.";
J. Clin. Microbiol. 37:3965-3970(1999).
EMBL: AJ132795; CAB6233.1;
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
Prodom: PD001149; Lipoprotein_6; 1.
SEQUENCE 207 AA; 22249 MW; 20899AA76E0C49D CRC64;

Query Match 31.9%; Score 878; DB 2; Length 207;
Best Local Similarity 98.9%; Pred. No. 6.2e-33;
Matches 180; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

379 NNSGSDASTNPDESAGKPNLTVISKITDTSNAFLAVKEVEALLSIDELSKAIGKKIK 438
DB 20 NNSGSDASTNPDESAGKPNLTVISKITDTSNAFLAVKEVEALLSIDELSKAIGKKIK 79
QY 439 NDGTLDNNEANRNESILAGAVEISKLTOKLSVLSNSELKKIKKEAKDCSOKFTTKLDSH 498
DB 80 NDGTLDNNEANRNESILAGAVEISKLTOKLSVLSNSELKKIKKEAKDCSOKFTTKLDSH 139
QY 499 AELGISVODDANKAKILKTHGTGDKGAKELFEKLSLESLSKAQAALNTSVKELTNPV 558
DB 140 AELGISVODDANKAKILKTHGTGDKGAKELFEKLSLESLSKAQAALNTSVKELTNPV 199
QY 559 VA 560
DB 200 VA 201

RESULT 14
Q44669 PRELIMINARY: PRT: 212 AA.
AC Q44669;
01-NOV-1996 (TREMBLrel. 01, Created)
01-NOV-1996 (TREMBLrel. 01, Last sequence update)
01-JUN-2001 (TREMBLrel. 17, Last annotation update)
OUTER SURFACE PROTEIN C.

```

```

GN OSPC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PRO;
RX MEDLINE=95395018; PubMed=7665660;
Jauris-Helpe S., Ilegi G., Preac-Mursic V., Roessler D., Schwab E.,
Rautschek E., Will G., Wilske B.;
"Molecular analysis of genes encoding outer surface protein C (OSPC)
of Borrelia burgdorferi sensu lato: relationship to ospa genotype and
evidence of lateral gene exchange of ospC.";
J. Clin. Microbiol. 33:1860-1866(1995).
EMBL: X81521; CA57241.1;
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
Prodom: PD001149; Lipoprotein_6; 1.
SEQUENCE 212 AA; 22514 MW; C6EDC33BFD488DDE CRC64;

Query Match 31.9%; Score 877; DB 2; Length 212;
Best Local Similarity 89.8%; Pred. No. 7.1e-33;
Matches 184; Conservative 4; Mismatches 3; Indels 14; Gaps 2;

QY 195 NNSGSDASTNPDESAGKPNLTVISKITDTSNAFLAVKEVEALLSIDELSKAIG 254
DB 20 NNSGSDASTNPDESAGKPNLTVISKITDTSNAFLAVKEVEALLSIDELSKAIG 79
QY 255 QKIDNNGLAALNNQNSLLAGAYASTLITKESKLNKEELTEITAKKCKSEPTNK 314
DB 80 QKIDNNGLAALNNQNSLLAGAYASTLITKESKLNKEELTEITAKKCKSEPTNK 139
QY 315 LKSHADLGGODADDEHAKAKILKTHATTDGKAEFDIESVGLKKAQVALTNSVKE 374
DB 140 LKSHADLGGODADDEHAKAKILKTHATTDGKAEFDIESVGLKKAQVALTNSVKE 199
QY 375 LGRNNSGSDASTNP--DESAGK 397
DB 200 L-----TSPIVVAESPKKP 212

RESULT 15
Q50623 PRELIMINARY: PRT: 203 AA.
AC Q50623;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PMUD;
RA Masuzawa T., Sawaki K., Yoshii T.;
Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
EMBL: AB009899; BAA24129.1;
InterPro: IPR001800; Lipoprotein_6.
Pfam: PF01441; Lipoprotein_6; 1.
Prodom: PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 203
FT NON_TER 1
SEQUENCE 203 AA; 21505 MW; 66DF78BBAE219 CRC64;

Query Match 31.9%; Score 876; DB 2; Length 203;
Best Local Similarity 98.9%; Pred. No. 7.5e-33;
Matches 179; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 195 NNSGSDASTNPDESAGKPNLTVISKITDTSNAFLAVKEVEALLSIDELSKAIG 254

```

```
Db      14  NNSGKGASASTNPADESAGBNLTEISKKITDSNAFVLAVKEVETVLVLSIDELAKKAIG 73
OY      255  OKIDNNGLAALNNONGSLAGATAISTLITEKLSKLNLEELKTEIAKAKKCEEFYTK 314
Db       74  OKIDNNGLAALNNONGSLAGATAISTLITEKLSKLNLEELKTEIAKAKKCEEFYTK 133
OY      315  LKSGHADLGKODATDDHAKAAILKTHATTDKGAKKEFKDLFESVEGLKKAQVALTNSYKE 374
Db      134  LKSGHADLGKODATDDHAKAAILKTHATTDKGAKKEFKDLFESVEGLKKAQVALTNSYKE 193
OY       375  L 375
Db      194  L 194
```

Search completed: March 18, 2002, 10:10:56
Job time: 979 sec

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☒ **OTHER:** page opposite of each other

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

